

QY 1125 CGGCTCCATCCGGCGCGTGGTCAACATCTTTCTGTGGGCAAGGGCGCTTGTCTCTATCC 1184
 Db 492 CGGTTCCATCCGGCGCGTGGTCAACATCTTCTGTGGGCAAGGGCGCTGTCTCTATCC 551
 QY 1185 TCTGCCATCTTTGGCGCTGTGAGGTGCTGGAGAGTGCCTTCCAGGAAGG 1238
 Db 552 GTTGCCCTCTTCTGGCGCGTGGAGTGTGGAGAGTCTCTTCTCCAGGAAGG 605

RESULT 15

AG576538/c

LOCUS

DEFINITION

Mus musculus molossinus DNA, clone:MSMg01-500M12.T7, genomic survey

sequence.

ACCESSION

AG576538

VERSION

AG576538.1

KEYWORDS

GSS.

ORGANISM

Mus musculus molossinus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.

BAC end Sequences of Library MSMg01

Unpublished

2 (bases 1 to 689)

Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.

Direct Submission

Submitted (17-NOV-2003) Mashira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan 305-0856, Japan

(E-mail: hattori@gs.c.riken.jp, URL: http://hgp.gs.c.riken.go.jp/)

Tel: 81-45-503-9111, Fax: 81-45-503-9170

Clones are derived from the mouse BAC library MSMg01. For BAC

library availability, please contact Kuniya Abe (abe@rtc.riken.jp).

Tsukuba Institute, Bio Resource Center,

The Institute of Physical and Chemical Research (RIKEN) 3-1-1

Koyadai, Tsukuba, 305-0074 Japan

Phone: 81-298-36-9189, fax: 81-298-36-9199

e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : T7

LIBRARY

Vector : pBAC3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI.

Location/Qualifiers

1..689

/organism="Mus musculus molossinus"

/mol_type="genomic DNA"

/sub_species="molossinus"

/db_xref="taxon:57486"

/clone="MSMg01-500M12.T7"

/sex="male"

/tissue type="mixture of kidney and spleen"

/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match

Best Local Similarity

Matches

574; Conservative

0; Mismatches

63; Indels

1; Gaps

1;

QY 388

CAGGGCATCTTCTGCTGGCGCTTACCTACCGCATCTGACGGCGGCTACCTGGGTTG 447

Db 653

CAGGGCATCTTCTGCTGGCGCTTACCTACCGCATCTGACGGCGGCTACCTGGGTTG 594

QY 448

TTTCTCATCATCTTCTGCTGGCGCTTACCTACCGCATCTGACGGCGGCTACCTGGGTTG 507

Db 593

TTTCTCATCATCTTCTGCTGGCGCTTACCTACCGCATCTGACGGCGGCTACCTGGGTTG 534

QY 508

CTGTACGAGAGATGAAGACGGCGAGGTGGTGGCGCTGGGAGTCTGTACGTGGCCATA 567

Db 533

CTGTACGAGAGATGAAGACGGCGAGGTGGTGGCGCTGGGAGTCTGTATGTGGCCATA 474

QY 568 GCCAACGCTGTGTGGCGCGCTTCCAAACGCTGGGGCGGCGAGTGGTGAACCTAGCG 627
 Db 473 GCTAACGATGCTGGCGCTTCCGATCCCGACCCCTGGGGCGGCGTGGTCAATGTGGCG 414
 QY 628 CAGATCATCGAGCTGGTGTGATGACGCTGATCTTGTACGCTGGTGGTGGTGAACCTCATG 687
 Db 413 CAGATCATCGAGCTGGTGTGATGACGCTGATCTTGTACGCTGGTGGTGGTGAACCTCATG 354
 QY 688 TACAACAGCTTCCCGGGGCTGCGGCTGTCGAGAGTCTGCTGCTTATATCGCCAGCGCC 747
 Db 353 TACAACAGCTTCCCGGGGCTGCGGCTGTCGAGAGTCTGCTGCTTATATCGCCAGCAA 294
 QY 748 GTGCTGCTGCTTGGCGCTTCTTAAGAACTCAAGGCGGCTGTCGAGTTCAGTCTGCTG 807
 Db 293 GTGCTGCTGCTTGGCGCTTCTTGAAGAATCTCAAGGCGGCTGTCGAGTTCAGTCTGCTG 234
 QY 808 TGCATCTGGGCGGCTTCTGCTCATCAATATCTGGTTCATAGCTTCTGCTATCGGGGCG 867
 Db 233 TGTACGCTGGGCGGCTTCTGCTCATCAACATCTGCTGCTTCTGCTTCTCTCTCTCTCC 174
 QY 868 CGCGACTGGGCGGCTGGGAGAGGTCAAGTCTTACATCGAGCTCAAGAGTTCCTCCGCTCC 927
 Db 173 CGTGATTTGGGCGGCTGGGAGAGGTGAAGTCTTACATCGAGCTCAAGAGTTCCTCCGCTCC 114
 QY 928 ATTGGCATCATCTGCTTTCAGCTTACAGCTTCTCAGATCTTCTGCTTCTGCTGGAGGCAAT 987
 Db 113 ATTGGCATCATCTGCTTTCAGCTTACAGCTTCTCAGATCTTCTGCTTCTGCTGGAGGCAAC 54
 QY 988 ATGCAGCAGCGGCGGCTTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1025
 Db 53 ATGCAGC-GCCGAGCGATTCCCGGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 17

Search completed: November 3, 2004, 22:22:57

Job time : 4621 secs

/db_xref="taxon:10090"
 /clone="IMAGE:6809167"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EVO"
 /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGGGTGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
 Query Match 34.1%; Score 537.4; DB 5; Length 689;
 Best Local Similarity 90.0%; Pred. No. 1.6e-106;
 Matches 611; Conservative 0; Mismatches 61; Indels 7; Gaps 3;

206 CTTGCGGGACGAGCGCTGAGCGCCCTCGAGGGAGACATCCATTATCAGGAGCA 265
 Db |||||
 14 CTTGCGGGACGAGCGCGAGAGCTCCCTCGAGGGAGACATTCATTATCAGC 70
 Qy |||||
 266 GCGGAGCTCTCTGCGCCCTCGCGCTCCAGGACACAG--GTGGGAGGTGGTGGCAAT 322
 Db |||||
 71 GCGGCGTCTCTGCGACCTCTGGCTCCAGGACAGCGCGTGGGAGTGGTGGGAGT 130
 Qy |||||
 323 TCGGGGCGCCACGAAACCCAAATACGCGGTGGGAGGAGCGTGGAACTGACCAACG 382
 Db |||||
 131 TCGGGGCGCCACGAAACCCAAATACGCGGTGGGAGGAGCGTGGAACTGACCAATG 190
 Qy |||||
 383 CCATCCAGGCGATGTTGCTGGGCTACCTACGCCATCCTGCACGGGGCTACTCG 442
 Db |||||
 191 CCATTACGGGCGATGTTGCTGGGCTACCTACGCCATCCTCCACGGGGCTACTCG 250
 Qy |||||
 443 GGTGTTTCTCATCATCTTCGCGCGCGCTGTGTGCTGTCTACACGGCAAGATCCTCATCG 502
 Db |||||
 251 GGTGTTTCTCATCATCTTCGCGCGAGTGTGTGCTGTCTACACGGCAAGATCCTCATCG 310
 Qy |||||
 503 CGTGCTGTACGAGGAGATGAGACGGCGAGTGTGTGCGCGTGGGAGCTCTGATCGTGG 562
 Db |||||
 311 CGTGCTGTACGAGGAGATGAGACGGCGAGTGTGTGCGCGTGGGAGCTCTGATGTGG 370
 Qy |||||
 563 CCATAGCCACGCTGTGTGCGCGCGCTTCCACAGCTT-GGGCGCGCGAGTGGTGAAC 621
 Db |||||
 371 CCATAGCTAACGATGTGTGCGCTTCTGATTCCTCCACCTTNGGCGCGCGTGGTCAAT 430
 Qy |||||
 622 GTAGCGCAGATCATCGAGCTGTGTGATGACGTGCATCTGTGATGTGTGTGAGTGCAAC 681
 Db |||||
 431 GTGGCGCAGATCATCGAGCTGTGTGATGACGTGTATCTTGTACGTGTGTGAGCGGCAAC 490
 Qy |||||
 682 CTCATGTACAAAGCTTCCCGGGGCTGCCCGTGTGCGAGAGTCTTGTTCCATTATCGCC 741
 Db |||||
 491 CTCATGTACAAAGTTCCTCCGGGGCTGCCCGTGTGCGAAAGTCTTGTTCCATATAGCC 550
 Qy |||||
 742 ACGCGCGTGTGTGCGCTTCCCGGCTTCTTAAGAACTTAAGAACTTAAGAACTTAAGAACT 801
 Db |||||
 551 ACGCGCGTGTGTGCGCTTCCCGGCTTCTTAAGAACTTAAGAACTTAAGAACTTAAGAACT 610
 Qy |||||
 802 CTGCTGTGCACTCTGGGCGCACTTCTGTCATCAATATCTGTGTATAGCTACTGTCTATCG 861
 Db |||||
 611 CTGCTGTGAGCTGTGGCGCACTTCTGTCATCAACATCTGTGTATCTGTCTACTGTCTCT 670
 Qy |||||
 862 CGGCGCGCGACTGGGCGCT 880

SOURCE
ORGANISM
Taeniopygia guttata
Taeniopygia guttata
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Passeriformes; Estrildidae;
Estrildinae; Taeniopygia.
REFERENCE
AUTHORS
1 (bases 1 to 805)
Clayton, D.F., Arnold, A.P., Ball, G.F., Brenowitz, E., George, J.M.,
Mello, C.V., Wade, J., Replogle, K., Lewin, H., Band, M., Hernandez, A.
and Liu, L.
The Songbird Neurogenomics Initiative: An Evolving Public Resource
for Study of Genes, Brain, and Behavior
Unpublished (2004)
CONTACT: David F. Clayton
University of Illinois
B107 CLSL, 601 S. Goodwin, Urbana, IL 61801, USA
Tel: 217 244 3668
Fax: 217 244 1648
Email: dclayton@uiuc.edu
Base Calling/Quality Scores: PHRED from Washington University
Genome Center
Vector Trimming: Cross_match from Washington University Genome
Center PHRAP suite. Low quality bases (Phred score < 20) were
trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length. Funded
by PHS grant # R01 NS045264, 'Songbird Neurogenomics Initiative.'
PCR Primers
FORWARD: TAATACGACTCACTATAGG(T7)
BACKWARD: ATTAAACCTCACTAAG(T3)
Insert Length: 805 Std Error: 0.00
Plate: SB02048A2 row: A column: 08
Seq primer: TAATACGACTCACTATAGG (T7)
High quality sequence stop: 805.
FEATURES
source
Location/Qualifiers
1..805
/organism="Taeniopygia guttata"
/mol_type="mRNA"
/db_xref="taxon:59729"
/clone="SB02048A2A08.f1"
/tissue_type="brain"
/dev_stage="late embryo, post-hatch days 1, 10, 20, 45,
and adult (pooled)"
/lab_host="DH10B"
/clone_lib="normalized Keck-Tagu Library SB02"
/note="Organ: brain; Vector: pBS II SK(+); Site 1:
EcoRI(5' side of insert); Site 2: NotI (3' side of
insert); The library was constructed and normalized as
described by Bonaldo, M.F., Lennon, G. and Soares, M.B.
(1996), Genome Research 6(9): 791-806. An identifying tag
was added at the 3' during cDNA synthesis:
insertAAAAAAAAAAAAAAAAAATGCGA."

ORIGIN
Query Match 34.4%; Score 542.6; DB 7; Length 805;
Best Local Similarity 79.4%; Pred. No. 1.2e-107;
Matches 638; Conservative 0; Mismatches 166; Indels 0; Gaps 0;
QY 630 GATCATCGAGCTGTCATGACGTCGATCTGTACGTGGTGGTCAAGCTCATGTA 689
DB 1 GATCATTTGAACCTGGTCATGACCTGCATCTCTATGTGGTGGTCAAGCTCATGTA 60
QY 690 CAACAGCTTCCCGGGGCTGCCGTGTGCGAGAACTCTGTGTCATTTATCGCCACGCCCT 749
DB 61 CAACAGCTTCCCGGCTGCTCCGAGAGTGTGGTTCATTTGCCACGGCACT 120
QY 750 GCTGCTGCCTTGGCCCTCTTAAAGAACTTCAAGGCCGTGTCAAGTTCAAGTCTGTGTG 809
DB 121 GCTCCTGCTTGGCGTCTTGAAGAACTTCAAGGCCGTGTCAAGTTCAAGTCTGTGTG 180
QY 810 CACTCTGGCCCACTTCTGTCATCAATATCTGGTCTAGCTACTGTCTATCGGGGCGG 869
DB 181 CAGTTTAGCCCACTTTGTTCATCAACATCTCGGTGATTTGCTCTCTCCAGGGCGG 240
QY 870 CGACTGGGCTGGGAGAGGTCAAGTTCTACATCGACGTCAAGAACTTCCCATCTCCAT 929

241 TGACTGGGCTGGGACAAAGTCAAGTTTACATGATGTCAAGAAGTTTCCCATCTCCAT 300
QY TGGCATCATCGTGTTCAGCTACAGCTCTCAGATCTTCTCGCTTCGCTGGAGGGCAATAT 989
DB TGGCATCATGTTCTTCAGCTACACCTCCAGATCTTTCGCTTCCTTCGAGGGGAACAT 360
QY GCAGCAGCCCAAGGAGTTCCACTGCATGATGAACCTGGACGACATCGCAGCTTCGCTGCT 1049
DB GCAGAACCCCAAGGAGTTTCATTCATGATGAACCTGGACTCACATAGCAGCTTGCATCT 420
QY 1050 CAAGGCTCTTTCGCGCTCGTTCGCTACCTCACCTGGGCCGACGAGACCAAGAGGTCTAT 1109
DB TAAGGGACTCTTTCGCTTCGCTACCTGACCTGCGCTGATGAGACCAAGGAGGTCTAT 480
QY 1110 CACGGATTAACCTGCGGCTCCATCCGCGCGTGGTCAACATCTTCTGTTGGCCCAAGGC 1169
DB TACAGACAACTTGCCATCCACCATTAGGGCAGTAGTCAATATTTCTTGGTGGCCAAAGC 540
QY 1170 GCTGTTGCTCTATCTCTGCGCTTCTTTCGCGCTGTCGAGGTGCTGGAGAACTCGCTCTT 1229
DB CTGCTCTCATACCCCTTTCGCTTCTTTCGAGCTGTAGAAGTCTCGGAGCGATCCCTTTT 600
QY 1230 CCAGAAAGGACGCGCGCTTTTTCGCGCTGCTTACAGCGGACGCGGCGCTGAAGTC 1289
DB CCAAGATGAAACAGGGCTTTCTTCCCAACTGCTATGCGGGGTGACGGGCGGCTCAAAATC 660
QY 1290 CTGGGGCTGACGCTGCGCTGCGCGCTGCTTTCACGCTGCTCATGGCCATTTATGT 1349
DB CTGGGGACTCACCTCTCAGATGTGCGCTTTCACCTGCTCATGGCTTANNATGT 720
QY 1350 GCGGCACCTTCGCGCTGCTCATGGGCTCACCGGCGCTCACGGGCGGCGCTCTGTTT 1409
DB CCCCATTTCCTCTGATGGGCTTACTGGGAGCTCACAGCGCANNCTCTGTTT 780
QY 1410 CTGCTGCCAGCTCTTTTTCACCT 1433
DB CTTGCTGCCAGCTCTCTTCCACCT 804
RESULT 13
LOCUS BU709443 689 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-EVO-cbg-e-06-0-UI.r1 NIH BMAP_EVO Mus musculus cDNA clone
IMAGE:6809167 5', mRNA sequence.
ACCESSION BU709443
VERSION BU709443.1 GI:23642916
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 689)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyx-5.
Location/Qualifiers
1..689
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"

FEATURES
source
1..689
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"

before RNA extraction and purification (Tri-reagent method). cDNA was primed using oligo-dT primer: 5'-TGACTAGTCTTAGATCGGAGCGCGCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb resulted in an average insert size of 2.18 kb. This primary library is not normalized (normalized library is NIH MGC 255) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH_MGC library"

ORIGIN

```
Query Match      37.4%; Score 590; DB 7; Length 763;
Best Local Similarity 91.1%; Pred. No. 5.4e-118;
Matches 638; Conservative 0; Mismatches 60; Indels 2; Gaps 1;

QY 702 GGGGTCGCCGTCGCGAGAGTCTGTGTCATTATCCGACGGCGGTCGTGTCGCTTG 761
DB 1 GGGGTCGCCGTCGCGAGAGTCTGTGTCATTATCCGACGGCGGTCGTGTCGCTTG 60

QY 762 CGCCTTCCTTAAGAACCTCAAGCGCGTGTCCAAAGTTCACTGTGTGTCACCTCTGCCCCA 821
DB 61 CGCCTTCCTGAAGATCTCAAGCGCGTGTCCAAAGTTCACTGTGTGTCACCTCTGCCCCA 120

QY 822 CTTCTGTCATCAATATCTTGTCATAGCCTACTGTCTATCGCGGCGGCGACCTGGGCTG 881
DB 121 CTTCTGTCATCAATATCTTGTCATAGCCTACTGTCTATCGCGGCGGCGACCTGGGCTG 180

QY 882 GAGAGAGTCAAGTTCTACATCGAGTCAAGAGTTCCCATCTCCATTCGATCGATCATG 941
DB 181 GAGAGAGTCAAGTTCTACATCGAGTCAAGAGTTCCCATCTCCATTCGATCGATCATG 240

QY 942 GTTTCAGTACAGTCTCAGATCTTCTGCTTCGCTGGAGGCAATATGACAGCCCCAG 1001
DB 241 GTTTCAGTACAGTCTCAGATCTTCTGCTTCGCTGGAGGCAATATGACAGCCCCAG 300

QY 1002 CGAGTTCCATCGATGAATGAGCGACATCGAGCGCTGCTGCTCAAGGGCTCTTT 1061
DB 301 CGAATTCATCGATGAATGAGCGACATCGAGCGCTGCTGCTCAAGGGCTCTTT 360

QY 1062 CGCGCTCGTCCGCTACCTCACTGGCGGCGAGACAGACAGGAGTCAATCAGGATACT 1121
DB 361 CGCGCTCGTCCGCTACCTCACTGGCGGCGAGACAGACAGGAGTCAATCAGGATACT 420

QY 1122 GCGCGCTCCATCCGCGCGTGTCAACATCTTCTGTGGCCAAAGCGCTGTGTCTTA 1181
DB 421 GCGCGCTCCATCCGCGCGTGTCAACATCTTCTGTGGCCAAAGCGCTGTGTCTTA 480

QY 1182 TCCTCTGCCATCTTTTGGCGCTGTGAGTGTGAGAGTCTCTTCCAGAAAGCGAG 1241
DB 481 CCGGTTGCGCTTCTTCCGCGCGTGTGAGTGTGAGAGTCTCTTCCAGAAAGCGAG 540

QY 1242 CGCGCGCTTTTCCGCGCGTGTGAGTGTGAGAGTCTCTTCCAGAAAGCGAG 1301
DB 541 TCGTGCCTTTTCCGCGCGTGTGAGTGTGAGAGTCTCTTCCAGAAAGCGAG 600

QY 1302 GCTGCGCTGCGGCTCGTCTTCACTGCTGTGAGTGTGAGAGTCTCTTCCAGAAAGCGAG 1361
DB 601 GCTGCGCTGCGGCTCGTCTTCACTGCTGTGAGTGTGAGAGTCTCTTCCAGAAAGCGAG 660

QY 1362 GTTGTCTCATGGGCTCACCGGAGCGCTCACGGGCGCGCGC 1401
DB 661 GCTGTCTCATGG--CCTCACGGGAGCTCACGGGAGCGGCG 698
```

RESULT 10

```
CN503909
LOCUS      22405600 NIH_ZGC_9 Danio rerio cDNA clone IMAGE:7273715
DEFINITION 5', mRNA sequence.
ACCESSION  CN503909
VERSION     CN503909.1 GI:46816533
KEYWORDS   EST.
SOURCE     Danio rerio (zebrafish)
ORGANISM   Danio rerio
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 937)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

AUTHORS

Contact: Daniela S. Gerhard, Ph.D.

JOURNAL

Office of Cancer Genomics

COMMENT

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabs-remail.nih.gov

Tissue Procurement: Chi-Bin Chien

cDNA Library Preparation: Dr. Sumio Sugano

cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

plate: LLAM15261 row: i column: 09

High quality sequence start: 8

High quality sequence stop: 698.

Location/Qualifiers

1..937

/organism="Danio rerio"

/mol_type="mRNA"

/db_xref="taxon:7955"

/clone="IMAGE:7273715"

/tissue_type="neural retina, retinal pigment epithelium,

lens and overlying skin, pooled embryos"

/lab_host="DH10B Tona"

/clone_lib="NIH_ZGC 9"

/note="Organ: eye; Vector: pME18S-FL3; Site 1: DraIII;

Site 2: DraIII; 1st strand cDNA was primed with an

oligo(dT) primer

[GCGCTGAAGACGCGCTATGCGCTTTTCTTTTCTTTT];

double-stranded cDNA was ligated to a DraIII adaptor

[GGCUACUGG], digested and directionally cloned into

distinct DraIII sites of the pME18S-FL3. Library was size

selected for 1.0 kb, with an average insert size of ~1.2kb,

and is not amplified. Library constructed by Yutaka Suzuki

(University of Tokyo Institute of Medical Science). Custom

primers recommended for sequencing: 5' end primer

5'-GGATGTTGCTTTACTTCTA-3' and 3' end primer

5'-CGACTGCGAGCTCGAGACA-3'. Note: This is a Zebrafish Gene

Collection (ZGC) library"

Collection (ZGC) library"

Query Match 36.9%; Score 582; DB 7; Length 937;

Best Local Similarity 81.2%; Pred. No. 3.1e-116;

Matches 675; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 468 CTTTGTGTGCTGCTACACGGGCAAGATCTCTATCGGTGCTGTAGGAGAGTAAGA 527

DB 6 CTTTGTGCTGCTACACGGGCAAGATCTCTATCGGTGCTGTAGGAGAGTAAGA 65

QY 528 CGGCGAGTGTGCGCGTGTGCGGAGTCTGCTACGTGCGCATAGCAACGCTGCTGCGCCCC 587

DB 66 CGGCGAGTGTGCGGAGTGTGAGAGTCTGCTACGTGAGACATCGCAACGCTGCTGCGCTC 125

QY 588 GCGCTTCCCAACGCTGCGGCGCGGAGTGTGAGACGTAGCGCAGATCATCGAGCTGTGTAT 647

DB 126 GCGCTTCCCGGCTCGGAGGACACGTCTGCTCAACGTGGCACAGATCATCGAGCTGTGTAT 185

QY 648 GAGCTGCTATCTGTAGTGTGCTGAGTGGCAACCTCATGTACACAGCTTCCCGGGCT 707

DB 186 GACCTGCTATTTTGTACGTGCTGTCAGCGGGAACCTGATGTATCAACAGCTTCCCCACCT 245

QY 708 GCGCGTGTGCGAAGAGTCTGCTGCTCATTTATCCGACGCGGCGGTGCTGCTGCTTGGCCCTT 767

DB 246 CCCAGTGTCCGAGAGTCTGCTGGGCGATCATCCGACCGCGGCTCTCTGCTTGGCCCTT 305

QY 768 CCTTAAGAACCTCAAGGCGCGTGTCCAAGTTCAGTCTGTGTGCTGCTGCGCCACCTTCGT 827

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 675)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..675
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30681555"
/tissue_type="whole eye"
/dev_stage="newborn (1, 5, 15 days)"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP HW0"
/notes="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tsg
sequence located between the Not I site and the polyA tail
is AATAATGAC. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."
37.6%; Score 593.4; DB 7; Length 675;
Best Local Similarity 92.4%; Pred. No. 9.7e-119;
Matches 624; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

ORIGIN

Query Match
Best Local Similarity 92.4%; Pred. No. 9.7e-119;
Matches 624; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 902 TCGACGTCAAGAGTTCCCATCTCCATTCGATCATCGTTGTCAGCTACAGTCTCAGA 961
DB 1 TCGACGTCAAGAGTTCCCATCTCCATTCGATCATCGTTGTCAGCTACAGTCTCAGA 60
QY 962 TCTTCTGCTTCGCTGGAGGGAATATGACAGAGCCAGCGAGTTCCAGTCCATCATGA 1021
DB 61 TCTTCTGCTTCGCTGGAGGGAATATGACAGAGCCAGCGAGTTCCAGTCCATCATGA 120
QY 1022 ACTGGACGCACATCGCAGCTGCTGCTCAAGGGCTCTTCGGGCTCGTCCGCTACCTCA 1081
DB 121 ACTGGACGCACATCGCAGCTGCTGCTCAAGGGCTCTTCGGGCTCGTCCGCTACCTCA 180
QY 1082 CTTGGCCGACGACAGACAGAGGTATCATCGATTAACCTGCCCGCTCATCCGGCGCG 1141
DB 181 CTTGGCCGACGACAGACAGAGGTATCATCGATTAACCTGCCCGCTCATCCGGCGCG 240
QY 1142 TGGTCAACATCTTCTTGGTGGCCAGCGCTGTGTCCTATCTCTCGGCTATCTTTCGCG 1201
DB 241 TGGTCAACATCTTCTTGGTGGCCAGCGCTGTGTCCTATCTCTCGGCTATCTTTCGCG 300
QY 1202 CTTGTCAGGTGCTGGAGAGTTCGCTCTTCAGGAGGCGAGCGCGCTTTTCCCGGCT 1261

DB 301 CCCTCGAAGTGTGGAGAAAGTCTCTCTCCAGGAAGGCAAGTCGCGGCTTCTTCCCGCCT 360
QY 1262 GCTACAGCGCGACGCGCGCTCAAGTCTCTGGGGGCTGACGCTGCGCTGCGCTCGTCG 1321
DB 361 GCTATGAGAGCGACGCGCTTAAAGTCTTGGGGGCTGACGCTGCGCTGCGCTGCTGG 420
QY 1322 TCTTCAAGCTGCTCATGGCCATTTATGTGCGGCACATTCGCGCTGCTCATGGGCTCACCG 1381
DB 421 TCTTCAAGCTGCTCATGGCCATTTATGTGCGGCACATTCGCGCTGCTCATGGGCTCACCG 480
QY 1382 GCAGCTCAAGCGCGCGCTCTGTTCTTGTGTCGCCAGCTCTTTCACCTGGGCTGTC 1441
DB 481 GCAGCTCAAGCGCGCGCTCTGTTCTTGTGTCGCCAGCTCTTTCACCTGGGCTGTC 540
QY 1442 TCTGGCGGACGCTGTCGCGACCAAGTCTTCTTCGACGTCGCGCATCTTTCGTCGGCG 1501
DB 541 TCTGGCGGACGCTGTCGCGACCAAGTCTTCTTCGATGTGGCCATCTTTCGTCATGGCG 600
QY 1502 GCATCTGCGAGCTGTCGCGCTTCGTCACCTCCCTCGAGGGCTCATCGAAGCTACCGAA 1561
DB 601 GCATCTGCGAGCTGTCGCGCTTCGTCACCTCATCGAAGCTCATCGAAGCTACCGAA 660
QY 1562 CCAACGCGAGGACT 1576
DB 661 CCAACGCGAGGACT 675

RESULT 9
CO393521 763 bp mRNA linear EST 01-JUL-2004
LOCUS
DEFINITION
IMAGE:7314034 5', mRNA sequence.
ACCESSION
CO393521
VERSION
CO393521.1 GI:49575437
KEYWORDS
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 763)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
College of Wisconsin
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM15366 row: a column: 08
High quality sequence stop: 738.
Location/Qualifiers
1..763
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7314034"
/sex="both"
/tissue_type="Brain - Pooled from several tissues from one
or more individuals"
/lab_host="DH10B TonA"
/clone_lib="NIH MGC 254"
/note="Organ: brain/CNS; Vector: pExpress-1; Site 1:
EcoRV; Site 2: NotI; RNA obtained from brain tissue of 8
wk old animal. Tissues were snap-frozen and kept at -80C


```
source
1. 778
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5717808"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FDO"
/notes="Organ: brain; Vector: pYX-Asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
```

ORIGIN

Query Match 40.4%; Score 638.2; DB 5; Length 778;
Best Local Similarity 90.3%; Pred. No. 1.7e-128;
Matches 705; Conservative 0; Mismatches 120; Indels 6; Gaps 2;

QY 103 GCGGCCACGATGAGGAGCGGTGGCTTCGCGCATTCGACGACCTCGACTTGGACAC 162
DB 1 GCGGCCACGATGAGGAGCGGTGGCTTCGCGCATTCGACGACCTCGACTTGGACAT 60
QY 163 CGCCAGGCGCTCGAGATGACATCTCTGAAAGCGGAGGAGCGCTCGCGGGACGAGGC 222
DB 61 CGCCAGGCGCTCGAGATGACATCTCTGAAATCGNAGGCGGCTCGGAGAGGCG 120
QY 223 GCTGAAGCGCGCTCGAGGAGACATCCATTATCAGGAGCGAGGAGCTCTTCGCG 282
DB 121 GCAGAGCTCCCGTCGAGGAGACATTCATTATCAGC--GCGGCGCGCTCTCTGCCA 177
QY 283 CCCTCGGCTCCAGGACAG--GTGGGAGGTGGTGGCAATTCGGGGCCACGACAG 339
DB 178 CCCTCGGCTCCAGGACAGCGCGCTGGAGCTGGTGGGAGTTGGGGGTCTACGACAAA 237
QY 340 CCCAAATACGCGCTGGGAGGAGCTGGAACGCTGACCAACGCAATCCAGGCGATGTT 399
DB 238 CCCAAGATCAGCGCTGGGAGGCGGCTGGAACTGACAAATGCCATTACAGGCGATGTT 297
QY 400 GTGCTGGGCTACCTACGCGCTTCGACGCGGCTACCTGGGGTGTTCATCATC 459
DB 298 GTGCTGGGCTACCTACGCGCTTCGACGCGGCTACCTGGGGTGTTCATCATC 357
QY 460 TTCGCGCGCTTGTGCTGTACACGCGCAAGATCTCATCGCGTCCCTGTACGAGGAG 519
DB 358 TTCGCGCGCTTGTGCTGTACACGCGCAAGATCTCATCGCGTCCCTGTACGAGGAG 417
QY 520 AATGAAGACCGCGAGGTGGTGGCGTGGGGAATCTCGTACGTGGCCATAGCAACGCTGC 579
DB 418 AACGAGACCGGAGGTGGTGGCGTGGGGAATCTGTATGTGGCCATAGTAACGATGC 477
QY 580 TGGCGCGCGCTTCCCAACGCTGGGCGCGGAGTGGTGAACCTAGCGCAGATCATCGAG 639
DB 478 TGGCGCTCTCGATTCCCAACCTTCGCGCGCGGCTGGTCAATGTGGCGCAGATCATCGAG 537
QY 640 CTGGTGATGACGTGCACTCTGTACGTGGTGGTGGCAACCTCATGTATACACAGCTTC 699
DB 538 CTGGTGATGACGTGATCTTGTACGTGGTGGTGGCAACCTCATGTATACACAGCTTC 597
QY 700 CCGGGGCTGCCCGTGCAGAAAGTCTCGTCCATTATCGCCACGCGCGTGCCTGCCT 759

598 CCGGGGCTGCCCGTGCAGAAAGTCTCGTCCATCATAGCCACAGCGGTGCTGCTGCC 657
760 TGGCGCTTCTTAAGACCTCAAGGCGGTGCTCAAGTTCAAGTTCAAGTTCAAGTTCA 819
658 TGGCGCTTCTTAAGAAATCTCAAGGCGGTGCTCAAGTTCAAGTTCAAGTTCAAGTT 717
820 CACTTCTCATCAATATCTTGGTCAATAGCTTCTATCTATCGGGGCGCGGACTGGGCC 879
718 CACTTCTCATCAATATCTTGGTCAATAGCTTCTATCTCTCTCGGGCGGTGATGGGCC 777
880 T 880
778 T 778

CD635843 658 bp mRNA linear EST 12-JAN-2004
5600431J1 FLP Homo sapiens cDNA, mRNA sequence.
CD635843
CD635843.1 GI:40284110
EST.
KEYWORDS
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 658)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS
Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL
Genomics 84 (1), 205-210 (2004)
COMMENT
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1. 658
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="PLP"
/note="Vector: pDrive Cloning Vector"

FEATURES
source
Query Match 38.5%; Score 607.8; DB 6; Length 658;
Best Local Similarity 99.1%; Pred. No. 7e-122;
Matches 642; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 443 GGTGTTTCTCATCATCTTCGCGCGGTGCTGTGCTGTACACGCGCAAGATCTCATCG 502
DB 658 GGTGTTTCTCATCATCTTCACGCGGTGCTGTGCTGTACACGCGCAAGATCTCATCG 599
QY 503 CGTGCCTGTACGAGGAGATGAAGACGCGGAGTGGTGGCGTGGCGGACTCGTACGTGG 562
DB 598 CGTGCCTGTACGAGGAGATGAAGACGCGGAGTGGTGGCGTGGCGGACTCGTACGTGG 539
QY 563 CCATAGCCAAACGCTCTCGCGCGCTTCCCAACGCTGGGCGCGCGAGTGGTGAACG 622
DB 538 CCATAGCCAAACGCTCTCGCGCGCTTCCCAACGCTGGGCGCGAGTGGTGAACG 480
QY 623 TAGCGCAGATCATCGAGCTGGTGATGACGTCGATCTCTGTACGTTGGTGGGCAACC 682
DB 479 TAGCGCAGATCATCGAGCTGGTGATGACGTCGATCTCTGTACGTTGGTGGGCAACC 420
QY 683 TCATGTACAAACAGCTTCCCGGGGCTCCCGTGTGCGAGAGTCTCTGTTCATTATCGCCA 742
DB 419 TCATGTACAAACAGCTTCCCGGGGCTCCCGTGTGCGAGAGTCTCTGTTCATTATCGCCA 360
QY 743 CGGCGCTGTGCTGCTTCGCGCTTCTTTAAGAACCTCAAGGCGGTGTCCAGTTTCAGTC 802
DB 359 CGGCGCTGTGCTGCTTCGCGCTTCTTTAAGAACCTCAAGGCGGTGTCCAGTTTCAGTC 300


```
QY 586 CCGCGCTTCCCAACGCTGGGCGCGAGTGTGAAGTGTAGCGAGATCATCGAGCTGGT 645
Db 181 CCGCGCTTCCCAACGCTGGGCGCGAGTGTGAAGTGTAGCGAGATCATCGAGCTGGT 240
QY 646 ATGACGTGCATCTCTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 705
Db 241 ATGACGTGCATCTCTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 706 CTGCGCGTGTGCGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 765
Db 301 TTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
QY 766 TTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 825
Db 361 TTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 826 GTCATCAATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 885
Db 421 GTCATCAATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY 886 AAGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 945
Db 481 AAGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 946 AGTGTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1005
Db 541 AGTGTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 1006 TTCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1065
Db 601 TTCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY 1066 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1125
Db 661 TTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY 1126 GGCTCCATCCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1185
Db 721 ACCACCATCCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 1186 CTGCGATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1245
Db 781 CTGCGATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
QY 1246 GCCTTTTCCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1305
Db 841 GCCTTTTCCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY 1306 CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1365
Db 901 AGGATTTCCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 958
QY 1366 CTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1425
Db 959 CTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1016
QY 1426 TTTTCC 1431
Db 1017 TTCCAC 1022
```

```
RESULT 3
BG700073
LOCUS 602681089F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4813726 5',
DEFINITION mRNA sequence.
ACCESSION BG700073
VERSION BG700073.1 GI:13969048
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 700)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10708 row: a column: 23
High quality sequence stop: 700.

FEATURES

source

1..700
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4813726"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_95"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.5 kb and normalized to 10^5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this
is a NIH_MGC Library."

ORIGIN

Query Match 43.0%; Score 677.8; DB 4; Length 700;
Best Local Similarity 98.9%; Pred. No. 3.9e-137;
Matches 693; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 146 ACCTCCACTTTGAGCACCAGGCGCTGAGATGACATCTCTGAAAGCCGAGGAGC 205
Db 1 AGCGGGGCTTTGAGCACCAGC -GGCTGAGATGACATCTCTGAAAGCCGAGGAGC 59
QY 206 CTGCGGGGACGAGGCGCTGAAGCCCGTCCAGGAGACATCCATTATCAGCAGGCA 265
Db 60 CTGCGGGGACGAGGCGCTGAAGCCCGTCCAGGAGACATCCATTATCAGCAGGCA 119
QY 266 GCGGAGCTCTCTGCGCGCTTCCAGGACCCAGGTGGAGGTGGTGGGAAATTCG 325
Db 120 GCGGAGCTCTCTGCGCGCTTCCAGGACCCAGGTGGAGGTGGTGGGAAATTCG 179
QY 326 GGGGCCACACAGCCAAAATCAGCGGTGGAGGAGGCTGGAACGTGACCAACGCCA 385
Db 180 GGGGCCACACAGCCAAAATCAGCGGTGGAGGAGGCTGGAACGTGACCAACGCCA 239
QY 386 TCCAGGGCATGTTGCTGGGCTACCTCTAGCCATCTGACGCGGCTACTGGGT 445
Db 240 TCCAGGGCATGTTGCTGGGCTACCTCTAGCCATCTGACGCGGCTACTGGGT 299
QY 446 TGTTCCTCATCTTTCGCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 505
Db 300 TGTTCCTCATCTTTCGCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 359
QY 506 GCCTGTACAGAGAGATGAAGACGCGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 565
Db 360 GCCTGTACAGAGAGATGAAGACGCGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 419
QY 566 TAGCCAAACGCTGT 625
Db 420 TAGCCAAACGCTGT 479

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 18:33:25 ; Search time 4611 Seconds
(without alignments)
12470.591 Million cell updates/sec

Title: US-09-940-919-1

Perfect score: 1578

Sequence: 1 atggccacctgtctgcgacg.....gaaccaacgaggactag 1578

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	849.2	53.8	1730	3 BC025013	Mus muscu
2	689.2	43.7	1027	9 CNS0292M	AL188727 Tetraodon
3	677.8	43.0	700	4 BG700073	BG700073 602681089
4	658.2	41.7	848	6 CD557601	CD557601 AGENCOURT
5	638.2	40.4	778	5 BQ573293	UI-M-PD0-BQ573293 UI-M-PD0-
6	607.8	38.5	658	6 CD635843	CD635843 56000431J
7	597.6	37.9	1031	9 CNS03DEA	AL238987 Tetraodon
8	593.4	37.6	675	7 CO429271	CO429271 UI-M-HW0-
9	590	37.4	763	7 CO393521	CO393521 AGENCOURT
10	582	36.9	937	7 CNS03909	CNS03909 AGENCOURT
11	544	34.5	624	7 CK619909	CK619909 mk29g12.Y
12	542.6	34.4	805	7 CK309005	CK309005 SB02048A2
13	537.4	34.1	689	5 BU709443	BU709443 UI-M-EV0-
14	525.2	33.3	605	6 CB581479	CB581479 AMGNNUC:N
15	525.2	33.3	689	9 AG576538	AG576538 Mus muscu
16	520.2	33.0	1101	9 CNS052RP	AL318526 Tetraodon
17	511.6	32.4	1080	9 CNS05L98	AL342485 Tetraodon
18	509	32.3	767	4 BJ735127	BJ735127 BJ735127
19	503.8	31.9	579	2 AW742427	up57a04.Y
20	495.4	31.4	595	8 AZ098746	RPCI-23-1
21	495	31.4	670	4 B1822066	B1822066 603039933
22	464.8	29.5	886	7 CN833492	AGENCOURT
23	464.4	29.4	549	8 AQ933231	RPCI-23-2
24	457.2	29.0	693	9 CNS04J95	AL293234 Tetraodon

25	456.6	28.9	863	7	CN838148	CN838148 AGENCOURT
26	444	28.1	508	2	BE981082	UI-M-CG0-BE981082 UI-M-CG0-
27	435.6	27.6	522	8	AZ002297	RPCI-23-3
28	433.4	27.5	751	4	BI488536	603021049
29	433	27.4	509	1	AI839048	UI-M-AKO-AI839048 UI-M-AKO-
30	429	27.2	608	4	BI393349	p9p1n.pk0-BI393349 p9p1n.pk0
31	419.2	26.6	480	2	BE982265	UI-M-CG0p-BE982265 UI-M-CG0p
32	410.2	26.0	471	2	AW120774	UI-M-BH2-AW120774 UI-M-BH2-
33	395	25.0	623	1	AL588280	AL588280 AL588280
34	393	24.9	487	1	AI836076	UI-M-AU0-AI836076 UI-M-AU0-
35	386.2	24.5	447	1	AI834859	UI-M-AN1-AI834859 UI-M-AN1-
36	382.2	24.2	734	5	BU339660	603406545
37	381.8	24.2	449	2	AW048723	UI-M-BH1-AW048723 UI-M-BH1-
38	367.2	23.3	867	7	CN833360	AGENCOURT
39	365.2	23.1	455	2	BE950161	UI-M-CE0-BE950161 UI-M-CE0-
40	362.6	23.0	456	1	AI837592	UI-M-AKO-AI837592 UI-M-AKO-
41	341.4	21.6	645	5	BU053668	UI-M-FD0-BU053668 UI-M-FD0-
42	337.4	21.4	457	7	H23502	Ym57h09.r1
43	334	21.2	942	6	CD361004	AGENCOURT
44	333	21.1	392	1	AI549592	UI-R-C3-t
45	328	20.8	403	2	BF543591	UI-R-C3-t

ALIGNMENTS

RESULT 1
BC025013/c
LOCUS BC025013 1730 bp mRNA linear HTC 07-AUG-2002
DEFINITION Mus musculus, clone IMAGE:5357287, mRNA.
ACCESSION BC025013
VERSION BC025013.1 GI:19035721
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1730)
Strausberg,R.
Direct Submission
Submitted (01-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: ang@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 54 Row: f Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis
This clone has the following problem: no polyA-tail.
Location/Qualifiers
1. 1730
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357287"
/tissue_type="Eye, retina, mouse strain C57Bl/6"

Db 1372 TGCTACAGCGCGACGGCGCGCTGAAGTCCTGGGGGCTGAGCTGCGCTGCGGCTCGTC 1431
Qy 1321 GTCTTACAGCTGCTCATGGCCATTATGTGCGCACTTTCGGCTGCTCATGGGCTCACC 1380
Db 1432 GTCTTACAGCTGCTCATGGCCATTATGTGCGCACCTTCGGCTGCTCATGGGCTCACC 1491
Qy 1381 GGCAGCCTCAGGGCGCGCGCTCTGTTTCTTGTGCCCCAGCCTCTTTTCACCTGCGGCTG 1440
Db 1492 GGCAGCCTCAGGGCGCGCGCTCTGTTTCTTGTGCCCCAGCCTCTTTTCACCTGCGGCTG 1551
Qy 1441 CTCTGGCGCAAGCTGCTGTGGCAACCAAGTCTTTCGACGTGCGCATCTTTCGTCATCGGC 1500
Db 1552 CTCTGGCGCAAGCTGCTGTGGCAACCAAGTCTTTCGACGTGCGCATCTTTCGTCATCGGC 1611
Qy 1501 GGCATCTGCAGCGTGTCCGGCTTGGTCACTCCCTCGAGGGCTCATCGAAGCTACCGA 1560
Db 1612 GGCATCTGCAGCGTGTCCGGCTTGGTCACTCCCTCGAGGGCTCATCGAAGCTACCGA 1671
Qy 1561 ACCAACGCGGAGGACTAG 1578
Db 1672 ACCAACGCGGAGGACTAG 1689

Search completed: November 3, 2004, 22:38:58
Job time : 788 secs

APPLICANT: Edinger, Shlomit
APPLICANT: Gerlach, Valerie
APPLICANT: Ellerman, Karen
APPLICANT: MacDougall, John R
APPLICANT: Gunther, Erik
APPLICANT: Millet, Isabelle
APPLICANT: Stone, David J
APPLICANT: Smithson, Glennda
APPLICANT: Szekeres Jr, Edward S
APPLICANT: Ji, Weizhen
TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding Th
TITLE OF INVENTION: Methods Of Using The Same
FILE REFERENCE: 21402-141 CIP
CURRENT APPLICATION NUMBER: US/10/096,625
PENDING FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 09/972,211
PRIOR FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 60/238,325
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 60/238,323
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 60/238,400
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,397
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,401
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,379
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,402
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 30/238,384
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,373
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,372
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,383
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,382
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/275,892
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/296,860
PRIOR FILING DATE: 2001-06-08
NUMBER OF SEQ ID NOS: 200
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 39
LENGTH: 1734
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (112)..(1686)
US-10-096-625-39

Query Match 99.8%; Score 1574.8; DB 16; Length 1734;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTGTCCGACGACGAGCTGTCCACGTCGCCACGTCCTCGTGTCCAAACAGTCC 60
DB 112 ATGGCCACCTTGTCCGACGACGAGCTGTCCACGTCGCCACGTCCTCGTGTCCAAACAGTCC 171
QY 61 CAGGCCAAGATGAGCGGCAATGTTCCGACGATGGGTTTTCAGCGGCCACCGATGAGGAG 120
DB 172 CAGGCCAAGATGAGCGGCAATGTTCCGACGATGGGTTTTCAGCGGCCACCGATGAGGAG 231
QY 121 CGGGTGGGCTTCGGCATGTCGACGACCTCGACTTTGACACCGCCAGGCGCTTGCAGATG 180
DB 232 CGGGTGGGCTTCGGCATGTCGACGACCTCGACTTTGACACCGCCAGGCGCTTGCAGATG 291
QY 181 GACATCCTGAAAGCCGAGGAGAGCCCTCGCGGGACGAGGGCGCTGAAAGCGCCCGTCGAG 240

292 GACATCCTGAAAGCCGAGGAGAGCCCTCGCGGGACGAGGGCGCTGAAAGCGCCCGTCGAG 351
241 GGAGACATCCATATCAGCGAGGAGCGGAGCTCCTCTGCGCGCCCTCGGCTCCAAAGAC 300
352 GGAGACATCCATATCAGCGAGGAGCGGAGCTCCTCTGCGCGCCCTCGGCTCCAAAGAC 411
301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCGACGACAAAGCCCAAAATCAGCGCGTGGAG 360
412 CAGGTGGAGGTGGTGGCGAATTCGGGGGCGACGACAAAGCCCAAAATCAGCGCGTGGAG 471
361 GCAGGCTGGAACGTGACCAACGCCATCAGGGCATGTCGTGCTGGGCGCTTACCTACGCC 420
472 GCAGGCTGGAACGTGACCAACGCCATCAGGGCATGTCGTGCTGGGCGCTTACCTACGCC 531
421 ATCTGTGACGCGGCTACCTGGGGTGTCTCTCATCATCTTCGCGCGCGTGTGTGTGC 480
532 ATCTGTGACGCGGCTACCTGGGGTGTCTCTCATCATCTTCGCGCGCGTGTGTGTGC 591
481 TACACCGGCAAGATCCTCATCGCTGCTGTACGAGGAGATGAAAGCGGCGAGTGTG 540
592 TACACCGGCAAGATCCTCATCGCTGCTGTACGAGGAGATGAAAGCGGCGAGTGTG 651
541 CGCGTGGGAGCTGTAGTGGCCATAGCCAAACGCTGCTGCGCCCGCTTCCCAACG 600
652 CGCGTGGGAGCTGTAGTGGCCATAGCCAAACGCTGCTGCGCCCGCTTCCCAACG 711
601 CTGGCGGCGGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGATGACGTGATCCTG 660
712 CTGGCGGCGGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGATGACGTGATCCTG 771
661 TACGTGTGTGAGTGGCAACCTCATGTATCAACAGTTCCTCGCGGCTGCGCGTGTGCGAG 720
772 TACGTGTGTGAGTGGCAACCTCATGTATCAACAGTTCCTCGCGGCTGCGCGTGTGCGAG 831
721 AAGTCTGTGTCATATCGCCACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
832 AAGTCTGTGTCATATCGCCACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
781 AAGCGGTGTCACAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
892 AAGCGGTGTCACAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951
841 GTCATAGCTACTGTCTATCGCGGCGCGAGCTGGGCTGCGGAGAGGTCAAGTTCATC 900
952 GTCATAGCTACTGTCTATCGCGGCGCGAGCTGGGCTGCGGAGAGGTCAAGTTCATC 1011
901 ATCGACGTCAAGAGTTCCTCCATTCGATGTCATGTCGTCATGTCGTCATGTCGTCATGTCG 960
1012 ATCGACGTCAAGAGTTCCTCCATTCGATGTCATGTCGTCATGTCGTCATGTCGTCATGTCG 1071
961 ATCTTCTGCTGCTGCTGCGAGGCAATATGACGAGCCAGCGAGTTCACCTGTCATGATG 1020
1072 ATCTTCTGCTGCTGCTGCGAGGCAATATGACGAGCCAGCGAGTTCACCTGTCATGATG 1131
1021 AACTGGACGACATCGACGCTGCTGCTCAAGGGCTTCTGCGCGCTGCTGCGCTACCTC 1080
1132 AACTGGACGACATCGACGCTGCTGCTCAAGGGCTTCTGCGCGCTGCTGCGCTACCTC 1191
1081 ACCTGGGCGGAGAGACCAAGAGGTGTCATGAGGATACCTGCGCGCTGCTGCGCTGCGCGC 1140
1192 ACCTGGGCGGAGAGACCAAGAGGTGTCATGAGGATACCTGCGCGCTGCTGCGCTGCGCGC 1251
1141 GTGGTCAACATCTTCTGTCGCGCAAGGCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1252 GTGGTCAACATCTTCTGTCGCGCAAGGCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1311
1201 GCTGTGAGAGTGTGTCGAGAGTGTGCTCTTCCAGGAGGAGCGCGCGCTTCTTCCGCGC 1260
1312 GCTGTGAGAGTGTGTCGAGAGTGTGCTCTTCCAGGAGGAGCGCGCGCTTCTTCCGCGC 1371
1261 TCGTACGCGGAGCGGCGCTGAGTCTGTCGCGGCTGACGCTGCGGCTGCGGCTGCGGCTGCTC 1320

;; PRIOR FILING DATE: 2000-10-06
;; PRIOR APPLICATION NUMBER: 60/275,892
;; PRIOR FILING DATE: 2001-03-14
;; PRIOR APPLICATION NUMBER: 60/296,860
;; PRIOR FILING DATE: 2001-06-08
;; NUMBER OF SEQ ID NOS: 200
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 37
;; LENGTH: 1734
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (112)..(1686)
US-10-096-625-37

Query Match 99.8%; Score 1574.8; DB 16; Length 1734;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTGTCTCCGACGAGCTGTCCAAAGTGGCCACGCTCCGTTGTCACCAAGTCC 60
DB 112 ATGGCCACCTTGTCTCCGACGAGCTGTCCAAAGTGGCCACGCTCCGTTGTCACCAAGTCC 171

QY 61 CAGGCCAAGATGAGCGGCATGTTCCGCCAGATGGGTTTTCAGGCGGCCACCGATGAGGAG 120
DB 172 CAGGCCAAGATGAGCGGCATGTTCCGCCAGATGGGTTTTCAGGCGGCCACCGATGAGGAG 231

QY 121 GCGGTGGGCTTCGCGCATTCGACGACCTCGACTTTGAGCACCGCCAGGSCCTTCGACATG 180
DB 232 GCGGTGGGCTTCGCGCATTCGACGACCTCGACTTTGAGCACCGCCAGGSCCTTCGACATG 291

QY 181 GACATCTGAAAGCCGAGGAGAGCCCTCGCGGAGCGAGGCGCTGAAGCGCCGCTCGAG 240
DB 292 GACATCTGAAAGCCGAGGAGAGCCCTCGCGGAGCGAGGCGCTGAAGCGCCGCTCGAG 351

QY 241 GGAGACATCATTTATCAGCGAGGCGAGCGAGCTCTCTGCGGCCCTCGGCTCCAAAGGAC 300
DB 352 GGAGACATCATTTATCAGCGAGGCGAGCGAGCTCTCTGCGGCCCTCGGCTCCAAAGGAC 411

QY 301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATTCAGCGCGTGGGAG 360
DB 412 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATTCAGCGCGTGGGAG 471

QY 361 GCAGGTGGAACTGACCAAGCCATCCAGGCGATGTTGCTGTGGGCTTACCTTAGCC 420
DB 472 GCAGGTGGAACTGACCAAGCCATCCAGGCGATGTTGCTGTGGGCTTACCTTAGCC 531

QY 421 ATCTGTCACGCGGCTACCTGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGCTGC 480
DB 532 ATCTGTCACGCGGCTACCTGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGCTGC 591

QY 481 TACACGGCAAGATCTCATCGGCTGCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 540
DB 592 TACACGGCAAGATCTCATCGGCTGCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 651

QY 541 CGCGTGGGACCTGACGTGGGCGATAGCAAGCCCTGTCGCGCCCGCGCTTCCCAAGC 600
DB 652 CGCGTGGGACCTGACGTGGGCGATAGCAAGCCCTGTCGCGCCCGCGCTTCCCAAGC 711

QY 601 CTGGGCGGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGAGTGCATCC 660
DB 712 CTGGGCGGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGAGTGCATCC 771

QY 661 TACGTGGTGGTGGACCTCATGTACAAAGCTTCCCGGGGCTGCCCGTGTGCGCAG 720
DB 772 TACGTGGTGGTGGACCTCATGTACAAAGCTTCCCGGGGCTGCCCGTGTGCGCAG 831

QY 721 AAGTCTGTGTCATTATCGCACGGCGGTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 780
DB 832 AAGTCTGTGTCATTATCGCACGGCGGTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 891

QY 781 AAGGCCGTGTCCAAAGTTTCAGTCTGCTGTGCACTCTCGGCCACTTCGTATCAATATCCTG 840

DB 892 AAGGCCGTGTCCAAAGTTTCAGTCTGCTGTGCACTCTGCCCCACTTCGTATCAATATCCTG 951
QY 841 GTCATAGCCTACTGTCTATCGCGGGCGCGCATGCGGCTTGGAGAAAGTCAAGTTCTAC 900
DB 952 GTCATAGCCTACTGTCTATCGCGGGCGCGCATGCGGCTTGGAGAAAGTCAAGTTCTAC 1011
QY 901 ATCAGCGTCAAGAAAGTTCCCATCTCCATTTGGGCATCATCGTGTTCAGCTACAGTCTCAG 960
DB 1012 ATCAGCGTCAAGAAAGTTCCCATCTCCATTTGGGCATCATCGTGTTCAGCTACAGTCTCAG 1071
QY 961 ATCTTCTGCTCTTCTGCTGAGGGAATATGACAGCGCCAGCGAGTTCCTACTGATGATG 1020
DB 1072 ATCTTCTGCTCTTCTGCTGAGGGAATATGACAGCGCCAGCGAGTTCCTACTGATGATG 1131
QY 1021 AACTGGAAGCAGCATCGCAGCCTCGTCTCAAGGGGCTCTTTCGCGCTCGTGCCTACCTC 1080
DB 1132 AACTGGAAGCAGCATCGCAGCCTCGTCTCAAGGGGCTCTTTCGCGCTCGTGCCTACCTC 1191
QY 1081 ACCTGGGCGCAGACGAGACCAAGAGGTCATCACGGATAAACCCTGCGGCTCCATCCGGGCC 1140
DB 1192 ACCTGGGCGCAGACGAGACCAAGAGGTCATCACGGATAAACCCTGCGGCTCCATCCGGGCC 1251
QY 1141 GTGCTCAACATCTTCTGCTGGCCAAAGGCGCTGTTGCTCTATCTCTGCGCATTTCTTTGCC 1200
DB 1252 GTGCTCAACATCTTCTGCTGGCCAAAGGCGCTGTTGCTCTATCTCTGCGCATTTCTTTGCC 1311
QY 1201 GCTGTCCAGGTGCTGGAGAGTCGCTCTTTCAGGAAAGGACGCGCGCTTTTTCGCGGCC 1260
DB 1312 GCTGTCCAGGTGCTGGAGAGTCGCTCTTTCAGGAAAGGACGCGCGCTTTTTCGCGGCC 1371
QY 1261 TGCTACAGCGGCGACCGGCGCTGAAAGTCTCTGGGGGCTGACGCTGCGCTGCGGCTGCTC 1320
DB 1372 TGCTACAGCGGCGACCGGCGCTGAAAGTCTCTGGGGGCTGACGCTGCGCTGCGGCTGCTC 1431
QY 1321 GTCTTCCAGCTGCTCATGCGCATTTATGTGCGGCATTTGCTGCGCTGCTCATGCGGCTCACC 1380
DB 1432 GTCTTCCAGCTGCTCATGCGCATTTATGTGCGGCATTTGCTGCGCTGCTCATGCGGCTCACC 1491
QY 1381 GGAGCCTCAAGCGGCGCGGCTCTGTTTCTTGTGCGCGAGCTCTTTCACCTGCGGCTG 1440
DB 1492 GGAGCCTCAAGCGGCGCGGCTCTGTTTCTTGTGCGCGAGCTCTTTCACCTGCGGCTG 1551
QY 1441 CTCTGGCGCAAGTGTGCTGGCACCAAGTCTTCTTCGAGCTGCGCATCTTCTGTCATCGGC 1500
DB 1552 CTCTGGCGCAAGTGTGCTGGCACCAAGTCTTCTTCGAGCTGCGCATCTTCTGTCATCGGC 1611
QY 1501 GGCACTCTGAGCGTGTCCGCGCTTCTGTCACCTCCCGAGGCTCATCGAAGCTTACCGA 1560
DB 1612 GGCACTCTGAGCGTGTCCGCGCTTCTGTCACCTCCCGAGGCTCATCGAAGCTTACCGA 1671
QY 1561 ACCAAGCGGAGGACTAG 1578
DB 1672 ACCAAGCGGAGGACTAG 1689

RESULT 15
US-10-096-625-39
; Sequence 39, Application US/10096625
; Publication No. US20040068095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li

301 CAGGTGGAGTGTGGCGAATTCGGGGGCGACAGCAAGCCCAAAATCACGGCGTGGAG 360
Db |||||
412 CAGGTGGAGTGTGGCGAATTCGGGGGCGACAGCAAGCCCAAAATCACGGCGTGGAG 471
Qy |||||
361 CGAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTGCTGCTGGGCTTACCTTACGCC 420
Db |||||
472 CGAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTGCTGCTGGGCTTACCTTACGCC 531
Qy |||||
421 ATCTTCCACGGCGGCTTACCTGGGCTGTTCTCATCATCTTCGCGCGCTGTTGCTGCTGC 480
Db |||||
532 ATCTTCCACGGCGGCTTACCTGGGCTGTTCTCATCATCTTCGCGCGCTGTTGCTGCTGC 591
Qy |||||
481 TACACGGCAAGATCTCATCCGCTGCTGTACGAGGAGAAATGAAGACGGCGAGTGGTG 540
Db |||||
592 TACACGGCAAGATCTCATCCGCTGCTGTACGAGGAGAAATGAAGACGGCGAGTGGTG 651
Qy |||||
541 CGCGTGGGCAATGCTAGCGCCATAGCCAAAGCGCTGTGCGCCCGCGGCTTCCCAAGC 600
Db |||||
652 CGCGTGGGCAATGCTAGCGCCATAGCCAAAGCGCTGTGCGCCCGCGGCTTCCCAAGC 711
Qy |||||
601 CTGGGCGGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGATGAGGTGCATCTCG 660
Db |||||
712 CTGGGCGGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGATGAGGTGCATCTCG 771
Qy |||||
661 TACGTGCTGCTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCCGCTGTGCGAG 720
Db |||||
772 TACGTGCTGCTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCCGCTGTGCGAG 831
Qy |||||
721 AAGTCTGCTGCTATATGCCCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db |||||
832 AAGTCTGCTGCTATATGCCCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891
Qy |||||
781 AAGGCGGCTGCTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db |||||
892 AAGGCGGCTGCTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 951
Qy |||||
841 GTCATAGCTTACTGCTATCGCGGGCGCGAGCTGGGCGCTGGGAGAGGTCAAGTTCTTAC 900
Db |||||
952 GTCATAGCTTACTGCTATCGCGGGCGCGAGCTGGGCGCTGGGAGAGGTCAAGTTCTTAC 1011
Qy |||||
901 ATGACGCTCAAGAAGTTCCTCATCTGCGATCATCGTGTGCTGCTGCTGCTGCTGCTGCTG 960
Db |||||
1012 ATGACGCTCAAGAAGTTCCTCATCTGCGATCATCGTGTGCTGCTGCTGCTGCTGCTGCTG 1071
Qy |||||
961 ATCTTCTGCTTCTGCTGAGGCAATATGACAGCCAGCGAGTTCACACTGCATGATG 1020
Db |||||
1072 ATCTTCTGCTTCTGCTGAGGCAATATGACAGCCAGCGAGTTCACACTGCATGATG 1131
Qy |||||
1021 AACTGACGACATCGCAGCTCGTGTCTCAAGGGCTCTTTCGCGCTCGTGCCTTACCTC 1080
Db |||||
1132 AACTGACGACATCGCAGCTCGTGTCTCAAGGGCTCTTTCGCGCTCGTGCCTTACCTC 1191
Qy |||||
1081 ACCTGGGCGCAGAGACCAAGAGGCTCATCCGGATAAAGCTGCGCGGCTCCATCCGCGCC 1140
Db |||||
1192 ACCTGGGCGCAGAGACCAAGAGGCTCATCCGGATAAAGCTGCGCGGCTCCATCCGCGCC 1251
Qy |||||
1141 GTGGTCAACATCTTCTGCTGAGGCAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db |||||
1252 GTGGTCAACATCTTCTGCTGAGGCAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1311
Qy |||||
1201 GCTGTGAGGTGCTGAGAGTGCCTTCTCAGAGAGGAGCGCGGCTTTTTCGCGGCG 1260
Db |||||
1312 GCTGTGAGGTGCTGAGAGTGCCTTCTCAGAGAGGAGCGCGGCTTTTTCGCGGCG 1371
Qy |||||
1261 TGCTACAGCGGCGAGCGGCGCTTGAAGTCTGCGGGGCTGACGCTGCGCTGCGCGCTGCTG 1320
Db |||||
1372 TGCTACAGCGGCGAGCGGCGCTTGAAGTCTGCGGGGCTGACGCTGCGCTGCGCGCTGCTG 1431
Qy |||||
1321 GTCTTACGCTGCTCATGGCAATATGTCGCGCACTTCCGCTGCTCATGGGCTTCAAC 1380
Db |||||
1432 GTCTTACGCTGCTCATGGCAATATGTCGCGCACTTCCGCTGCTCATGGGCTTCAAC 1491
Qy |||||
1381 GGCAGCCTCACGGGCGCGGCTCTGTTTCTGTGTCGCCAGCGCTCTTTTACCTGCGGCTG 1440

RESULT 14

US-10-096-625-37
; Sequence 37, Application US/10096625
; Publication No. US20040068095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jk, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennda
; APPLICANT: Szekeres Jr, Edward S
; APPLICANT: Ji, Weizhen
; TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141 CIP
; CURRENT APPLICATION NUMBER: US/10/096,625
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/972,211
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382

Db |||||
1492 GGCAGCCTCACGGGCGCGGCTCTGTCTTGTGTCGCCAGCCTCTTTACCTGCGCCTG 1551
Qy |||||
1441 CTCTGGCGCAGCTGTGTGACCAAGAGTCTTCTTGCAGCTGCGCATCTTCTGTCATCGGC 1500
Db |||||
1552 CTCTGGCGCAGCTGTGTGACCAAGAGTCTTCTTGCAGCTGCGCATCTTCTGTCATCGGC 1611
Qy |||||
1501 GGCATCTGCAGCTGTGTGCGGCTTCTGTCACATCTCCCTGAGGGCCTCATCGAAGCCTACCGA 1560
Db |||||
1612 GGCATCTGCAGCTGTGTGCGGCTTCTGTCACATCTCCCTGAGGGCCTCATCGAAGCCTACCGA 1671
Qy |||||
1561 ACCAAGCGGAGGACTAG 1578
Db |||||
1672 ACCAAGCGGAGGACTAG 1689

QY 1501 GGCACTCTGAGGCTGCGGCTTCGGTGCACCTCCCTCGAGGGGCTCATCGAAGCTACCGA 1560
Db |||||||
QY 1561 ACCAACGGGAGGACTAG 1578
Db |||||||

RESULT 12
US-10-096-625-31
; Sequence 31, Application US/10096625
; Publication No. US20040068095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glenda
; APPLICANT: Szekeres Jr, Edward S
; APPLICANT: Ji, Weizhen
; TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141 CIP
; CURRENT APPLICATION NUMBER: US/10/096, 625
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/972,211
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 31
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (112)..(1686)
US-10-096-625-31

Query Match 99.8%; Score 1574.8; DB 16; Length 1734;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCACTTGTCTCCGAGCAAGCTGTCCAAACGTGGCCAGCTCCGTGTCCAAACAGTCC 60
Db |||||||
QY 61 CAGGCCAAGATGAGCGGCATGTTGCCAGGATGGGTTTTTCAGGGGCCACCGATAGGAG 120
Db |||||||
QY 172 CAGGCCAAGATGAGCGGCATGTTGCCAGGATGGGTTTTTCAGGGGCCACCGATAGGAG 231
QY 121 GCGGTGGGCTTCGGGCATTGCGACCTGCACTTTGAGCACCGCCAGGGGCGCTGAAAGCGCCCTCGAG 180
Db |||||||
QY 232 GCGGTGGGCTTCGGGCATTGCGACCTGCACTTTGAGCACCGCCAGGGGCGCTGCAAGT 291
QY 181 GACATCTGAAACCGAGCGAGAGCCCTGCGGGGACGAGGGCGCTGAAAGCGCCCTCGAG 240
Db |||||||
QY 292 GACATCTGAAACCGAGCGAGAGCCCTGCGGGGACGAGGGCGCTGAAAGCGCCCTCGAG 351
QY 241 GGAGACATCCATTATACGAGGAGCGAGAGCTCTCTGCGGCGCTTCGCGGTCCAAAGGAC 300
Db |||||||
QY 352 GGAGACATCCATTATACGAGGAGCGAGAGCTCTCTGCGGCGCTTCGCGGTCCAAAGGAC 411
QY 301 CAGGTGGGAGGTGGTGGGGAATTCGGGGGCCACGACCAAGCCCAAAATCACGGCGTGGGAG 360
Db |||||||
QY 412 CAGGTGGGAGGTGGTGGGGAATTCGGGGGCCACGACCAAGCCCAAAATCACGGCGTGGGAG 471
QY 361 CAGGAGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGCTGGGCGCTACCTACGCC 420
Db |||||||
QY 472 GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGCTGGGCGCTACCTACGCC 531
QY 421 ATCTTGCACGGCGGTACTCTGGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGTGCTGC 480
Db |||||||
QY 532 ATCTTGCACGGCGGTACTCTGGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGTGCTGC 591
QY 481 TACACGGGCAGATCCTCATCGCGTGCCTGTACAGGAGAGATGAAGACGGCGAGGTGGT 540
Db |||||||
QY 592 TACACGGGCAGATCCTCATCGCGTGCCTGTACAGGAGAGATGAAGACGGCGAGGTGGT 651
QY 541 CCGGTGCGGGACTCGTACGTGGCCATAGCCACGCTGTGCGCGCGCGCTTCCCAACG 600
Db |||||||
QY 652 CCGGTGCGGGACTCGTACGTGGCCATAGCCACGCTGTGCGCGCGCGCTTCCCAACG 711
QY 601 CTGGGCGGCGGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGAATGACGTGCAATCCTG 660
Db |||||||
QY 712 CTGGGCGGCGGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGAATGACGTGCAATCCTG 771
QY 661 TACGTGGTGGTGAAGTGGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCGTGCGCAG 720
Db |||||||
QY 772 TACGTGGTGGTGAAGTGGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCGTGCGCAG 831
QY 721 AAGTCTGTGTCATTATCGCCACGCGCTGTGCTGCTTTCCTTAAAGAACCTC 780
Db |||||||
QY 832 AAGTCTGTGTCATTATCGCCACGCGCTGTGCTGCTTTCCTTAAAGAACCTC 891
QY 781 AAGGCGGTGTCCAAAGTTCAGTCTGTGTGCACTCTGGGCCCACTTCGTGCAATCAATATCCTG 840
Db |||||||
QY 892 AAGGCGGTGTCCAAAGTTCAGTCTGTGTGCACTCTGGGCCCACTTCGTGCAATCAATATCCTG 951
QY 841 GTCATAGCCTACTGTCTATCGCGGGGCGCGAGCTGGGCTGGGAGAGGTCAAGTTCTAC 900
Db |||||||
QY 952 GTCATAGCCTACTGTCTATCGCGGGGCGCGAGCTGGGCTGGGAGAGGTCAAGTTCTAC 1011

FILE REFERENCE: 21402-141
CURRENT APPLICATION NUMBER: US/09/972,211
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 60/238,325
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 60/238,323
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 60/238,400
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,397
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,401
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,379
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,402
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 30/238,384
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,373
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,372
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,383
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,382
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/275,892
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/296,860
PRIOR FILING DATE: 2001-06-08
NUMBER OF SEQ ID NOS: 198
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 39
LENGTH: 1734
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (112)..(1686)
US-09-972-211-39

Query Match 99.8%; Score 1574.8; DB 11; Length 1734;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGCCACCTTGCTCCGAGCAAGCTGTCCAACTGGCCAGCTCCGTTGCCAACCAAGTCC 60
Db 112 ATGCCACCTTGCTCCGAGCAAGCTGTCCAACTGGCCAGCTCCGTTGCCAACCAAGTCC 171
Qy 61 CAGGCCAAGATGAGCGCATGTTCCGCAAGATGGGTTTCAGGCGGCCACGGATGAGGAG 120
Db 172 CAGGCCAAGATGAGCGCATGTTCCGCAAGATGGGTTTCAGGCGGCCACGGATGAGGAG 231
Qy 121 GCGGTGGCTTCGCGCATTTGGACGACCTCGACTTTGAGCACGCCAGGGGCTCCAGATG 180
Db 232 GCGGTGGCTTCGCGCATTTGGACGACCTCGACTTTGAGCACGCCAGGGGCTCCAGATG 291
Qy 181 GACATCTGAAAGCCGAGGGAGACCTCTGCGGGACGAGGCGCTGAAGCGCCCGTCGAG 240
Db 292 GACATCTGAAAGCCGAGGGAGACCTCTGCGGGACGAGGCGCTGAAGCGCCCGTCGAG 351
Qy 241 GGACACATCCATTATCAGCGAGGACGGAGCTCTCTGCGGCCCTCCGGCTCCAAGGAC 300
Db 352 GGACACATCCATTATCAGCGAGGACGGAGCTCTCTGCGGCCCTCCGGCTCCAAGGAC 411
Qy 301 CAGGTGGAGGTGGTGGCGAATTCGGGGCCACGACAAAGCCAAAATCAGGCGTGGGAG 360
Db 412 CAGGTGGAGGTGGTGGCGAATTCGGGGCCACGACAAAGCCAAAATCAGGCGTGGGAG 471
Qy 361 GCAGGCTGGACGTCGACCAACGCCATCCAGGGCATGTTCTGCTGGGCGCTACCCCTAGCC 420
Db 472 GCAGGCTGGACGTCGACCAACGCCATCCAGGGCATGTTCTGCTGGGCGCTACCCCTAGCC 531

Qy 421 ATCTGTGACCGCGGCTACCTGGGGTTCTTCTCATCATCTTCGCGCGGCTTGTGTGCTGC 480
Db 532 ATCTGTGACCGCGGCTACCTGGGGTTCTTCTCATCATCTTCGCGCGGCTTGTGTGCTGC 591
Qy 481 TACACCGGCAAGATCCTCATCGCGTGCCTGTACGAGGAGAAATGAAGACGGCGAGTGGT 540
Db 592 TACACCGGCAAGATCCTCATCGCGTGCCTGTACGAGGAGAAATGAAGACGGCGAGTGGT 651
Qy 541 CGGTGCGGGAATCGTAAGTGGCCATAGCAACGCTGTGCGCGCCCGCGCTTCCCAACG 600
Db 652 CGGTGCGGGAATCGTAAGTGGCCATAGCAACGCTGTGCGCGCCCGCGCTTCCCAACG 711
Qy 601 CTGGCGCGGAGTGGTGAAGTAGCCAGATCATCGAGCTGTGATGACGTGATCCTG 660
Db 712 CTGGCGCGGAGTGGTGAAGTAGCCAGATCATCGAGCTGTGATGACGTGATCCTG 771
Qy 661 TACGTGTGTGTAGTGGCAACCTCATGTACAAAGCTTCCCGGGCTGCGCGTGTGCGAG 720
Db 772 TACGTGTGTGTAGTGGCAACCTCATGTACAAAGCTTCCCGGGCTGCGCGTGTGCGAG 831
Qy 721 AAGTCTGTGTCCATTATCGCAAGCGCGTGTGCTGCTGCTTGTGCGCTTCTTAAAGAACTC 780
Db 832 AAGTCTGTGTCCATTATCGCAAGCGCGTGTGCTGCTGCTTGTGCGCTTCTTAAAGAACTC 891
Qy 781 AAGCGCGTGTCCAAAGTTCAGTCTGCTGTGCACTCTGCGCCCACTTCGTCACTCAATATCCTG 840
Db 892 AAGCGCGTGTCCAAAGTTCAGTCTGCTGTGCACTCTGCGCCCACTTCGTCACTCAATATCCTG 951
Qy 841 GTCATAGCTTACTGTCTATCGCGGGCGCGAGCTGGGCTTGGGAGAAAGGTCAAGTTCCTAC 900
Db 952 GTCATAGCTTACTGTCTATCGCGGGCGCGAGCTGGGCTTGGGAGAAAGGTCAAGTTCCTAC 1011
Qy 901 ATCGACGTCAAGAAAGTTCCTCATTTGGCATCATCGTGTTCAGTCACTACAGTCTCAG 960
Db 1012 ATCGACGTCAAGAAAGTTCCTCATTTGGCATCATCGTGTTCAGTCACTACAGTCTCAG 1071
Qy 961 ATCTTCTGCTTCTGCTGGAGGGCAATATGACGAGCGCCAGCGAGTTCGATGATGATG 1020
Db 1072 ATCTTCTGCTTCTGCTGGAGGGCAATATGACGAGCGCCAGCGAGTTCGATGATGATG 1131
Qy 1021 AACTGGACGACATCGCAGCTGCTGCTCAAGGGCTCTTCGCGCTCGCTCGCTACCTC 1080
Db 1132 AACTGGACGACATCGCAGCTGCTGCTCAAGGGCTCTTCGCGCTCGCTCGCTACCTC 1191
Qy 1081 ACTTGGGCGCAAGAGGAGTCAATCAAGGATTAACCTGCGCGGCTCCATCCCGGCTC 1140
Db 1192 ACTTGGGCGCAAGAGGAGTCAATCAAGGATTAACCTGCGCGGCTCCATCCCGGCTC 1251
Qy 1141 GTGGTCAACATCTTCTGCTGGGCAAGGGCTGTTGCTCTATCTCTGCGCATTTCTTGGC 1200
Db 1252 GTGGTCAACATCTTCTGCTGGGCAAGGGCTGTTGCTCTATCTCTGCGCATTTCTTGGC 1311
Qy 1201 GCTGTGAGGTGCTGAGAAAGTCTCTTCCAGAAAGGAGCGCGCGCTTTTCCCGGCTC 1260
Db 1312 GCTGTGAGGTGCTGAGAAAGTCTCTTCCAGAAAGGAGCGCGCGCTTTTCCCGGCTC 1371
Qy 1261 TGCTACAGCGGCGACGGGCGCTGAAAGTCTGCGGGGCTGACGCTGCGCTCGCGCTCGT 1320
Db 1372 TGCTACAGCGGCGACGGGCGCTGAAAGTCTGCGGGGCTGACGCTGCGCTCGCGCTCGT 1431
Qy 1321 GTCTTCAAGCTGCTATGGCCATTATGTCGCGCATTTGCGCGTGTCTATGGGCGCTCAC 1380
Db 1432 GTCTTCAAGCTGCTATGGCCATTATGTCGCGCATTTGCGCGTGTCTATGGGCGCTCAC 1491
Qy 1381 GGCAGCTTCAAGCGCGCGGCTCTGTCTTCTGCTGCGCGCGCTTTCACCTGCGGCTG 1440
Db 1492 GGCAGCTTCAAGCGCGCGGCTCTGTCTTCTGCTGCGCGCGCTTTCACCTGCGGCTG 1551
Qy 1441 CTCTGCGGCAAGCTGTGTGGCAACCAAGTCTTCTGCAAGTCTGCGCGCTTTCGCTCATCG 1500
Db 1552 CTCTGCGGCAAGCTGTGTGGCAACCAAGTCTTCTGCAAGTCTTCTGCGCGCTTTCGCTCATCG 1611

FEATURE:
; NAME/KEY: CDS
; LOCATION: (112)...(1686)
US-09-972-211-37

Query Match 99.8%; Score 1574.8; DB 11; Length 1734;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCACTTCTCGGAGCAAGCTGTCCAAAGTGGCCACGTCGTGTCCAAAGTCC 60
DB 112 ATGGCCACCTTCTCGGAGCAAGCTGTCCAAAGTGGCCACGTCGTGTCCAAAGTCC 171
QY 61 CAGGCCAAGATGAGGGGATGTTCCGAGGATGGTTTTCCAGCGGCGCACGGATCAGGAG 120
DB 172 CAGGCCAAGATGAGGGGATGTTCCGAGGATGGTTTTCCAGCGGCGCACGGATCAGGAG 231
QY 121 GCGGTGGGCTTCGCGCATTTGCGACGACCTCGACTTTGAGCACCGCCAGGGCCTGCAGATG 180
DB 232 GCGGTGGGCTTCGCGCATTTGCGACGACCTCGACTTTGAGCACCGCCAGGGCCTGCAGATG 291
QY 181 GACATCTGAAGCCGAGGGAGAGCCCTCGGGGAGAGGGCGCTGAAGCGCCGCTCGAG 240
DB 292 GACATCTGAAGCCGAGGGAGAGCCCTCGGGGAGAGGGCGCTGAAGCGCCGCTCGAG 351
QY 241 GGAGACATCATATCAGCGAGCGAGCGAGCTCTCTGCGCGCCCTCGGGCTCCAAGGAC 300
DB 352 GGAGACATCATATCAGCGAGCGAGCGAGCTCTCTGCGCGCCCTCGGGCTCCAAGGAC 411
QY 301 CAGGTGGGAGGTGGTGGCGAAATTCGGGGGCGCACGACCAAGCCCAAAATCACGGCGTGGAG 360
DB 412 CAGGTGGGAGGTGGTGGCGAAATTCGGGGGCGCACGACCAAGCCCAAAATCACGGCGTGGAG 471
QY 361 GCAGGTGGAACGTGACCAACGCGCATCCAGGGCATGTTGCTGCTGGGCCCTACCTACGCC 420
DB 472 GCAGGTGGAACGTGACCAACGCGCATCCAGGGCATGTTGCTGCTGGGCCCTACCTACGCC 531
QY 421 ATCTCTGACGGCGGTACCTGGGTTGTTTCTCATCATCTTCCGCGCGTGTGTGCTGC 480
DB 532 ATCTCTGACGGCGGTACCTGGGTTGTTTCTCATCATCTTCCGCGCGTGTGTGCTGC 591
QY 481 TACACGGCAAGATCCTCATCGGTGCTGTATACGAGGAGAAATGAAGACGGCGAGGTGGT 540
DB 592 TACACGGCAAGATCCTCATCGGTGCTGTATACGAGGAGAAATGAAGACGGCGAGGTGGT 651
QY 541 CCGCTCGGGACTCGTATCGTGGCATAGCCAGCGCTGTGCGCCCGCGCTTCCAAAG 600
DB 652 CCGCTCGGGACTCGTATCGTGGCATAGCCAGCGCTGTGCGCCCGCGCTTCCAAAG 711
QY 601 CTGGCGCGCGAGTGGTGAACGTAGCGAGATCATCGAGTGGTATGACGTGCATCTCG 660
DB 712 CTGGCGCGCGAGTGGTGAACGTAGCGAGATCATCGAGTGGTATGACGTGCATCTCG 771
QY 661 TAGCTGTGTGAGTGGCAACCTCATGTATCAACAGCTTCCCGGGGTGCGCGTGTGCGAG 720
DB 772 TAGCTGTGTGAGTGGCAACCTCATGTATCAACAGCTTCCCGGGGTGCGCGTGTGCGAG 831
QY 721 AAGTCTGTGTCCATTATCGCCAGCGCGGTGCTGCTGCTGCTGCTGCTTAAAGACCTC 780
DB 832 AAGTCTGTGTCCATTATCGCCAGCGCGGTGCTGCTGCTGCTGCTTAAAGACCTC 891
QY 781 AAGCGCGTGTCCAAAGTTCAGTCTGTGTCGACTCTGCGCCACTTCTGATCATATCTCG 840
DB 892 AAGCGCGTGTCCAAAGTTCAGTCTGTGTCGACTCTGCGCCACTTCTGATCATATCTCG 951
QY 841 GTCATAGCTTACTGTCTATCGGGGCGCGAGCTGGGCGCTGGGAGAGGTCAAGTCTTAC 900
DB 952 GTCATAGCTTACTGTCTATCGGGGCGCGAGCTGGGCGCTGGGAGAGGTCAAGTCTTAC 1011
QY 901 ATCAGCGTCAAGAAGTTCCCATCTCCATTGGCATCATCGTGTTCAGCTACAGCTCTAG 960
DB 1012 ATCAGCGTCAAGAAGTTCCCATCTCCATTGGCATCATCGTGTTCAGCTACAGCTCTAG 1071

RESULT 11
US-09-972-211-39
; Sequence 39, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosbe, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glenn
; APPLICANT: Szekeles Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides Encoding T
; TITLE OF INVENTION: Methods Of Using The Same

QY 421 ATCTGTGACGGCGGTACCTGGGTTGTTTCTCATCATCTTTCGCCGCGTGTGTGTCG 480
Db 532 ATCTGTGACGGCGGTACCTGGGTTGTTTCTCATCATCTTTCGCCGCGTGTGTGTCG 591
QY 481 TACACCGCAGATCCTCATCGCTGCTGTACAGGAGATGAAGACGGCGAGGTGGT 540
Db 592 TACACCGCAGATCCTCATCGCTGCTGTACAGGAGATGAAGACGGCGAGGTGGT 651
QY 541 CGCGTGGGGACTGTAGTGGCCATAGCAACGCCCTGCTGCGCCCGCGCTTCCCAACG 600
Db 652 CGCGTGGGGACTGTAGTGGCCATAGCAACGCCCTGCTGCGCCCGCGCTTCCCAACG 711
QY 601 CTGGGCGCCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGACGTGATCCTG 660
Db 712 CTGGGCGCCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGACGTGATCCTG 771
QY 661 TACGTGGTGTGAGTGGACCTCATGTACACAGCTTCCGGGGCTGCCGCTGTGCGAG 720
Db 772 TACGTGGTGTGAGTGGACCTCATGTACACAGCTTCCGGGGCTGCCGCTGTGCGAG 831
QY 721 AAGTCTGTGTCATTATCGCACGGCGTGTGCTGCTTGGCGCTTCCCTTAAGAACCTC 780
Db 832 AAGTCTGTGTCATTATCGCACGGCGTGTGCTGCTTGGCGCTTCCCTTAAGAACCTC 891
QY 781 AAGGCGGTGTCCAAAGTTCAAGTCTGCTGTGCACTCTGGCCCACTTCGTTCATCAATATCCTG 840
Db 892 AAGGCGGTGTCCAAAGTTCAAGTCTGCTGTGCACTCTGGCCCACTTCGTTCATCAATATCCTG 951
QY 841 GTCATAGCTACTCTATCGCGGGCGCGACTTGGGCTTGGGAGAGGTCAAGTTCTAC 900
Db 952 GTCATAGCTACTCTATCGCGGGCGCGACTTGGGCTTGGGAGAGGTCAAGTTCTAC 1011
QY 901 ATGACCTCAAGAGTTCCCATCTCCATTCGGCATCATGCTGTTACGTACACGTCTCAG 960
Db 1012 ATGACCTCAAGAGTTCCCATCTCCATTCGGCATCATGCTGTTACGTACACGTCTCAG 1071
QY 961 ATCTTCTGCTTCCGCTGGAGGGCAATATGACGAGCCGACGAGTTCCTACGTATGATG 1020
Db 1072 ATCTTCTGCTTCCGCTGGAGGGCAATATGACGAGCCGACGAGTTCCTACGTATGATG 1131
QY 1021 AACTGGAAGCATCGCAGCTGCTCTCAAGGGCTTTCGCGCTCGTCCGCTACCTC 1080
Db 1132 AACTGGAAGCATCGCAGCTGCTCTCAAGGGCTTTCGCGCTCGTCCGCTACCTC 1191
QY 1081 ACCTGGCCGACGAGACCAAGAGGTATCACGGATACTGCCCCGCTCCATCCGGCC 1140
Db 1192 ACCTGGCCGACGAGACCAAGAGGTATCACGGATACTGCCCCGCTCCATCCGGCC 1251
QY 1141 GTGTCAACATCTTCTGTGGCCCAAGCGCTGTGTCTTATCTCTGCGCATTTCTTGCC 1200
Db 1252 GTGTCAACATCTTCTGTGGCCCAAGCGCTGTGTCTTATCTCTGCGCATTTCTTGCC 1311
QY 1201 GCTGTGAGTGTGGAGAGTGCCTCTTCCAGGAGCGCGCGCTTTTCCCGGCC 1260
Db 1312 GCTGTGAGTGTGGAGAGTGCCTCTTCCAGGAGCGCGCGCTTTTCCCGGCC 1371
QY 1261 TGCTACAGCGGACGCGGCGCTGAAGTCTGGGGCTGACGCTGCGCTGCGGCTGCTC 1320
Db 1372 TGCTACAGCGGACGCGGCGCTGAAGTCTGGGGCTGACGCTGCGGCTGCGGCTGCTC 1431
QY 1321 GTCTTACGCTGTATGGCCATTATGTGCCGACCTTCGCGTGTCTATGGGCTCACC 1380
Db 1432 GTCTTACGCTGTATGGCCATTATGTGCCGACCTTCGCGTGTCTATGGGCTCACC 1491
QY 1381 GGCAGCTTACGGGCGCGGCTCTGTGTTCTGTGCGCAGGCTTCTTACCTGCGGCTG 1440
Db 1492 GGCAGCTTACGGGCGCGGCTCTGTGTTCTGTGCGCAGGCTTCTTACCTGCGGCTG 1551
QY 1441 CTCTGGCGAAGCTGTGTGGCAACAGTCTTTCGACGTGCGCATCTTCGTATCGGC 1500
Db 1552 CTCTGGCGAAGCTGTGTGGCAACAGTCTTTCGACGTGCGCATCTTCGTATCGGC 1611
QY 1501 GGCATCTGCAGCGTGTCCGCGTTCGTGCACTCTCCCTCGAGGGCTCATCGAAGCCTACCGA 1560

Db 1612 GGCATCTGCAGCGTGTCCGCGTTCGTGCACTCCCTCGAGGGCCTCATCGAAGCCTACCGA 1671
QY 1561 ACCAACGCGGAGGACTAG 1578
Db 1672 ACCAACGCGGAGGACTAG 1689
RESULT 10
US-09-972-211-37
; Sequence 37, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennnda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1 Human Proteins, Polynucleotides Encoding T
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo sapiens

```
Db 1072 ATCTTCCTGCTTCTGCTGAGGGCAATATGAGCAGACCCAGGAGTTCACCTGATGATG 1131
Qy 1021 AACTGAGCAGCATCGCAGCTGCTGCTCAAGGGCTTTCGCGCTGCTGCTTACCTC 1080
Db 1132 AACTGAGCAGCATCGCAGCTGCTGCTCAAGGGCTTTCGCGCTGCTGCTTACCTC 1191
Qy 1081 ACCTGGGCGGAGCAGACCAAGAGGTATCATCGGATAACCTGCGGGCTCCATCGCGCC 1140
Db 1192 ACCTGGGCGGAGCAGACCAAGAGGCTCATCGGATAACCTGCGGGCTCCATCGCGCC 1251
Qy 1141 GTGGTCAACATCTTTCTGGTGCCCAAGGCTGTTGTCTATCTCTGCTTCTGCTTCTG 1200
Db 1252 GTGGTCAACATCTTTCTGGTGCCCAAGGCTGTTGTCTATCTCTGCTTCTGCTTCTG 1311
Qy 1201 GCTGTCAGGCTGCTGAGAGTCTCTTTCAGAGAGGAGCGCGGCTTTTCCGCGCC 1260
Db 1312 GCTGTCAGGCTGCTGAGAGTCTCTTTCAGAGAGGAGCGCGGCTTTTCCGCGCC 1371
Qy 1261 TGCTACAGCGGAGCGGCGCTGAAGTCTCTGGGGCTGAGCGTCTGCGCTGCGCTGCTC 1320
Db 1372 TGCTACAGCGGAGCGGCGCTGAAGTCTCTGGGGCTGAGCGTCTGCGCTGCGCTGCTC 1431
Qy 1321 GTCTTCAGGCTCTCATGCGCAATTTATGTGCGGCACTTTCGCGCTGCTCATGCGCTCAC 1380
Db 1432 GTCTTCAGGCTCTCATGCGCAATTTATGTGCGGCACTTTCGCGCTGCTCATGCGCTCAC 1491
Qy 1381 GCGAGCTTCAGGGGCGCGGCTCTGTTCTGCTGCGGAGCTTTCACCTGCGGCTG 1440
Db 1492 GCGAGCTTCAGGGGCGCGGCTCTGTTCTGCTGCGGAGCTTTCACCTGCGGCTG 1551
Qy 1441 CTCTGCGCAAGCTGCTGCGGAGCAAGTCTTCTGAGCTGCGCATCTTCTGATCGGC 1500
Db 1552 CTCTGCGCAAGCTGCTGCGGAGCAAGTCTTCTGAGCTGCGCATCTTCTGATCGGC 1611
Qy 1501 GGCATCTGCGAGGCTGCGGCTTCTGCTGCACTCCCTCGAGGGCTCATGAGCCTACCGA 1560
Db 1612 GGCATCTGCGAGGCTGCGGCTTCTGCTGCACTCCCTCGAGGGCTCATGAGCCTACCGA 1671
Qy 1561 ACCAAGCGGAGGACTAG 1578
Db 1672 ACCAAGCGGAGGACTAG 1689

RESULT 9
US-09-972-211-35
; Sequence 35, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malvankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook Jr, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glenda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1 Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141
```

```
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112)..(1686)
US-09-972-211-35

Query Match 99.8%; Score 1574.8; DB 11; Length 1734;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCACCTTGTCTCCGAGCAAGCTGTCCAAAGTGGCGACGTCGTGTCACCAAGTCC 60
Db 112 ATGGCCACCTTGTCTCCGAGCAAGCTGTCCAAAGTGGCGACGTCGTGTCACCAAGTCC 171
Qy 61 CAGGCCAAGATGAGCGGATGTTCCGAGGATGGGTTTTCAGGGCGGCCACCGATGAGGAG 120
Db 172 CAGGCCAAGATGAGCGGATGTTCCGAGGATGGGTTTTCAGGGCGGCCACCGATGAGGAG 231
Qy 121 GCGGTGGGCTTCGCGCATTTGGACACCTCGACTTTTGAGACCGCCAGGCGCTGCAGATG 180
Db 232 GCGGTGGGCTTCGCGCATTTGGACACCTCGACTTTTGAGACCGCCAGGCGCTGCAGATG 291
Qy 181 GACATCTTGAAGCGGAGGAGCCCTGCGGGGACGAGGCGCTGAGCGCGCCGCTCGAG 240
Db 292 GACATCTTGAAGCGGAGGAGCCCTGCGGGGACGAGGCGCTGAGCGCGCCGCTCGAG 351
Qy 241 GGAGACATCCATTATCAGCGGACGAGCGAGCTCTCTGCGCGCCCTTCCGCTCCAAGGAC 300
Db 352 GGAGACATCCATTATCAGCGGACGAGCGAGCTCTCTGCGCGCCCTTCCGCTCCAAGGAC 411
Qy 301 CAGGTGGGAGGTGGTGGCGAATTCCGGGGCGCACGACAGCCCAAAATCAGCGGCTGGGAG 360
Db 412 CAGGTGGGAGGTGGTGGCGAATTCCGGGGCGCACGACAGCCCAAAATCAGCGGCTGGGAG 471
Qy 361 CAGGCTGAGAGTCAACCAAGCCATCCAGGGGATGTTGCTGCTGGGCTTACCTACGCC 420
Db 472 CAGGCTGAGAGTCAACCAAGCCATCCAGGGGATGTTGCTGCTGGGCTTACCTACGCC 531
```

Db 1600 GGCACTCTGCGGCTTGGTCACTCCCTCGAGGCGCTCATCGAAGCCCTACCGA 1659

Qy 1561 ACCAACGGGAGGACTAG 1578
Db 1660 ACCAACGGGAGGACTAG 1677

RESULT 8

US-09-972-211-31
; Sequence 31, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zernusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides Encoding TH
; FILE REFERENCE: 21402-141
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: CDS
; LOCATION: (112)..(1686)
US-09-972-211-31

Query Match 99.8%; Score 1574.8; DB 11; Length 1734;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCACCTTGCTCCGACGAAAGCTGTCCAAAGTGGCCACGTCGCTGTCACAAAGTCC 60
Db 112 ATGGCCACCTTGCTCCGACGAAAGCTGTCCAAAGTGGCCACGTCGCTGTCACAAAGTCC 171
Qy 61 CAGSCCAAGATGAGCGCATGTTCCGACAGATGGTTCAGGGCGGCCAGGATGAGGAG 120
Db 172 CAGSCCAAGATGAGCGCATGTTCCGACAGATGGTTCAGGGCGGCCAGGATGAGGAG 231
Qy 121 GCGGTGGGCTTCGCGCATTCGACGACCTCGACTTTTGGAGCAGCCAGGAGGCTTCAGATG 180
Db 232 GCGGTGGGCTTCGCGCATTCGACGACCTCGACTTTTGGAGCAGCCAGGAGGCTTCAGATG 291
Qy 181 GACATCTGAAAGCGGAGGAGAGCCCTGCGGGAGCGAGGGCGCTGAAGCGCCGTCGAG 240
Db 292 GACATCTGAAAGCGGAGGAGAGCCCTGCGGGAGCGAGGGCGCTGAAGCGCCGTCGAG 351
Qy 241 GGAGACATCATTTATCAGCGAGGAGCGGAGCTCTTCCGCCCTCCGCTCCAGGAC 300
Db 352 GGAGACATCATTTATCAGCGAGGAGCGGAGCTCTTCCGCCCTCCGCTCCAGGAC 411
Qy 301 CAGGTGGAGGTGCTGGCGAATTCGGGGGCGACGACAAAGCCCAAAATCAGGGGTGGGAG 360
Db 412 CAGGTGGAGGTGCTGGCGAATTCGGGGGCGACGACAAAGCCCAAAATCAGGGGTGGGAG 471
Qy 361 CGAGGTGGAACGTGACCAACCGCATTCAGGGGAGTGTGCTGCTGGGCTTACCCCTAGCC 420
Db 472 CGAGGTGGAACGTGACCAACCGCATTCAGGGGAGTGTGCTGCTGGGCTTACCCCTAGCC 531
Qy 421 ATCCTGACGGCGGCTACCTGGGGTGTTCCTCATCATCTTCCGCCGCTGCTGCTGC 480
Db 532 ATCCTGACGGCGGCTACCTGGGGTGTTCCTCATCATCTTCCGCCGCTGCTGCTGC 591
Qy 481 TACACCGCAAGATCCTCATTCGCGTGCCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 540
Db 592 TACACCGCAAGATCCTCATTCGCGTGCCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 651
Qy 541 CGCGTGGGAGCTCGTACGTGGCCATAGCCAAAGCTGTGCGCCCGCGCTTCCCAACG 600
Db 652 CGCGTGGGAGCTCGTACGTGGCCATAGCCAAAGCTGTGCGCCCGCGCTTCCCAACG 711
Qy 601 CTGGGCGCGGAGTGTGAGCGTAGCGAGATCATCGAGCTGGTGTGATGAGCTGCATCCTG 660
Db 712 CTGGGCGCGGAGTGTGAGCGTAGCGAGATCATCGAGCTGGTGTGATGAGCTGCATCCTG 771
Qy 661 TACGTGCTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCCCGTGTGCGAG 720
Db 772 TACGTGCTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCCCGTGTGCGAG 831
Qy 721 AAGTCTGTGTCATTTATCGCCACGCGGCTGTGCTGCTGCGCTTCCCTTAAGAACCTC 780
Db 832 AAGTCTGTGTCATTTATCGCCACGCGGCTGTGCTGCTGCGCTTCCCTTAAGAACCTC 891
Qy 781 AAGGCGGTGTCCAAGTTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 892 AAGGCGGTGTCCAAGTTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 951
Qy 841 GTCATAGCCTACTGTCTATCGGGGCGCGACTGGGCGCTGGGAGAGGTCAAAGTTCTAC 900
Db 952 GTCATAGCCTACTGTCTATCGGGGCGCGACTGGGCGCTGGGAGAGGTCAAAGTTCTAC 1011
Qy 901 ATCGAGTCAAGAAGTTCCCATCTCCATTTGGCATCATGCTGTTTCAAGTACAGTCTCAG 960
Db 1012 ATCGAGTCAAGAAGTTCCCATCTCCATTTGGCATCATGCTGTTTCAAGTACAGTCTCAG 1071
Qy 961 ATCTTCTGCTTTCGCTGGAGGGCAATATGACAGCCCGAGGCTTCCACTGCATGATG 1020

Db 1461 GTCTTACAGCTGCTCATGCGCAATTTATGTGCGCGCACTTTCGCGCTGCTCATGCGGCTCACC 1520
Qy 1381 GCAGGCTTACGGCGCGGCTCTGTTTCTGCTGCCAGCCTCTTTCACCTGCGCCTG 1440
Db 1521 GGCAGCCTTACGGCGCGGCTCTGTTTCTGCTGCCAGCCTCTTTCACCTGCGCCTG 1580
Qy 1441 CTCTGGCGCAAGCTGCTGTGGCACCAAGTCTTCTTCGAGCTGCGCATCTTCTGTCATCGGC 1500
Db 1581 CTCTGGCGCAAGCTGCTGTGGCACCAAGTCTTCTTCGAGCTGCGCATCTTCTGTCATCGGC 1640
Qy 1501 GGCATCTGAGCGTGTGCGGCTTCTGTCATCTCCCTCGAGGGCTCATCGAAGCCTACCGA 1560
Db 1641 GGCATCTGAGCGTGTGCGGCTTCTGTCATCTCCCTCGAGGGCTCATCGAAGCCTACCGA 1700

Qy 1561 ACCAAGCGGAGACTAG 1578

Db 1701 ACCAAGCGGAGACTAG 1718

RESULT 7

US-10-239-566-1

; Sequence 1, Application US/10239566

; Publication No. US20030069398A1

; GENERAL INFORMATION:

; APPLICANT: Merck Patent GmbH

; TITLE OF INVENTION: New human GABA receptor

; FILE REFERENCE: VGATFRWS

; CURRENT APPLICATION NUMBER: US/10/239,566

; CURRENT FILING DATE: 2002-09-24

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 1800

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (100)..(1674)

US-10-239-566-1

Query Match

Best Local Similarity 99.9%; Score 1576.4; DB 14; Length 1800;

Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCCACCTTGTCTCGCGAGCAAGCTGTCCAGTGGCGCACGTCGCTGTCACCAAGTCC 60
Db 100 ATGGCCACCTTGTCTCGCGAGCAAGCTGTCCAGTGGCGCACGTCGCTGTCACCAAGTCC 159

Qy 61 CAGGCCAAGATGAGCGGCATGTTTCGCGAGGATGGGTTTTTCAGCGCGCCACGGATGAGGAG 120
Db 160 CAGGCCAAGATGAGCGGCATGTTTCGCGAGGATGGGTTTTTCAGCGCGCCACGGATGAGGAG 219

Qy 121 GCGGTGGGCTTCGCGCATTTGCGACGACCTGCACTTTGAGCACCGCGCAGGCTTCGACATG 180
Db 220 GCGGTGGGCTTCGCGCATTTGCGACGACCTGCACTTTGAGCACCGCGCAGGCTTCGACATG 279

Qy 181 GACATCTGAAGCCCGAGGAGACCTCGCGGAGAGGGCGCTCAAGCGCCGCTCGAG 240
Db 280 GACATCTGAAGCCCGAGGAGACCTCGCGGAGAGGGCGCTCAAGCGCCGCTCGAG 339

Qy 241 GGAGACATCCATTTATCAGCGAGGAGCGGAGCTCTTCGCGCCCTCCGCGCTCCAAAGGAC 300
Db 340 GGAGACATCCATTTATCAGCGAGGAGCGGAGCTCTTCGCGCCCTCCGCGCTCCAAAGGAC 399

Qy 301 CAGGTGGGAGTGGTGGCGAATTCGGGGGCCACGAAAGCCCAAAATTCAGCGGCTGGGAG 360
Db 400 CAGGTGGGAGTGGTGGCGAATTCGGGGGCCACGAAAGCCCAAAATTCAGCGGCTGGGAG 459

Qy 361 GCAGGTGGAAAGCTGACCAAGCCATCCAGGCGATGTTCTGCTGGGCTTACCTACGCC 420
Db 460 GCAGGTGGAAAGCTGACCAAGCCATCCAGGCGATGTTCTGCTGGGCTTACCTACGCC 519

Qy 421 ATCTGACGGCGGCTACCTGGGGTTGTTTCTCATCATCTTCGCGCGGCTGTGTGCTGC 480

Db 520 ATCTGACGGCGGCTACCTGGGGTTGTTTCTCATCATCTTCGCGCGGCTGTGTGTGCTGC 579
Qy 481 TACACCGCAGATCTCTCATCGCTGCTGTACGAGGAGAAATGAAGACGGCGAGGTGGTG 540
Db 580 TACACCGCAGATCTCTCATCGCTGCTGTGTACGAGGAGAAATGAAGACGGCGAGGTGGTG 639
Qy 541 CCGGTGGGGGACTCGTAGTGGCCATAGCCAAACGCTGCTGCGCCCGCTTCCCAAG 600
Db 640 CCGGTGGGGGACTCGTAGTGGCCATAGCCAAACGCTGCTGCGCCCGCTTCCCAAG 699
Qy 601 CTGGCGGCGCAGTGGTGAACGTAGCCGAGATCATCGAGCTGGTATGACGTGATCTCTG 660
Db 700 CTGGCGGCGCAGTGGTGAACGTAGCCGAGATCATCGAGCTGGTATGACGTGATCTCTG 759
Qy 661 TACGTGGTGGTGAAGTGGCAACCTCATGTACAAACAGCTTCCCGGGGCTCCCGTGTGCGAG 720
Db 760 TACGTGGTGGTGAAGTGGCAACCTCATGTACAAACAGCTTCCCGGGGCTCCCGTGTGCGAG 819
Qy 721 AAGTCTCGTCCATTTATCGCCACGCGCTGTGTGCTTGGCTTCTTAAAGAACCTC 780
Db 820 AAGTCTCGTCCATTTATCGCCACGCGCTGTGTGCTTGGCTTCTTAAAGAACCTC 879
Qy 781 AAGCCGTGTCCAAAGTTCAGTCTGCTGTGCACTTGGGCCAATTCGTCTCATCAATATCTG 840
Db 880 AAGCCGTGTCCAAAGTTCAGTCTGCTGTGCACTTGGGCCAATTCGTCTCATCAATATCTG 939
Qy 841 GTCTAGTACTGTCTATCGCGCGCGCACTTGGGCTTGGGAGAAAGTCAAGTCTTCTAC 900
Db 940 GTCTAGTACTGTCTATCGCGCGCGCACTTGGGCTTGGGAGAAAGTCAAGTCTTCTAC 999
Qy 901 ATCGAGCTCAAGAGTTCCTCATTCATTTGGCATCATCTGTTTTCAGCTTACAGCTCTCAG 960
Db 1000 ATCGAGCTCAAGAGTTCCTCATTCATTTGGCATCATCTGTTTTCAGCTTACAGCTCTCAA 1059
Qy 961 ATCTTCTGCTTCTGCTGAGGGCAATATGAGAGCCGAGGAGTTCACCTGCAATGATG 1020
Db 1060 ATCTTCTGCTTCTGCTGAGGGCAATATGAGAGCCGAGGAGTTCACCTGCAATGATG 1119
Qy 1021 AACTGAGCGCATCATCGAGCCTGCTGCTCAAGGGCCTTTCGCGCTGCTCGCCTACCTC 1080
Db 1120 AACTGAGCGCATCATCGAGCCTGCTGCTCAAGGGCCTTTCGCGCTGCTCGCCTACCTC 1179
Qy 1081 ACCTGGGCGGACGAGACCAAGAGGTTCATACGATAACCTGCCGGCTTCATTCGCGGC 1140
Db 1180 ACCTGGGCGGACGAGACCAAGAGGTTCATACGATAACCTGCCGGCTTCATTCGCGGC 1239
Qy 1141 GTGGTCAACATCTTCTGCTGCGCAAGGGCTGTTGCTCTATCTCTGCGCATCTTTCG 1200
Db 1240 GTGGTCAACATCTTCTGCTGCGCAAGGGCTGTTGCTCTATCTCTGCGCATCTTTCG 1299
Qy 1201 GCTGTGAGGTGCTGAGAGTCTCTTTCAGAGGAGGAGCGCGGCTTTTTTCCGGGCT 1260
Db 1300 GCTGTGAGGTGCTGAGAGTCTCTTTCAGAGGAGGAGCGCGGCTTTTTTCCGGGCT 1359
Qy 1261 TGCTACAGCGGCGACGGCGCTGAAAGTCTGCGGGGCTGACGCTGCGCTGCGCTGCTG 1320
Db 1360 TGCTACAGCGGCGACGGCGCTGAAAGTCTGCGGGGCTGACGCTGCGCTGCGCTGCTG 1419
Qy 1321 GTCTTCAAGCTGCTCATGCGCATTTATGTGCGGCACTTTCGCGCTGCTCATGGGCTTACC 1380
Db 1420 GTCTTCAAGCTGCTCATGCGCATTTATGTGCGGCACTTTCGCGCTGCTCATGGGCTTACC 1479
Qy 1381 GGCAGCTCACGGGCGCGGCTCTGTTTCTGCTGCCAGGCTCTTTTCACTGCGGCTG 1440
Db 1480 GGCAGCTCACGGGCGCGGCTCTGTTTCTGCTGCCAGGCTCTTTTCACTGCGGCTG 1539
Qy 1441 CTCTGCGCAAGCTGCTGTGGCACCAAGTCTTCTTCGAGCTGCGGCACTTTCGCTCATCGGC 1500
Db 1540 CTCTGCGCAAGCTGCTGTGGCACCAAGTCTTCTTCGAGCTGCGGCACTTTCGCTCATCGGC 1599
Qy 1501 GGCATCTGAGCGTGTGCGCTTGTGTCATCTCCCTCGAGGGCTCATCGAAGCCTACCGA 1560


```
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennnda
; APPLICANT: Szekeres Jr, Edward S
; APPLICANT: Ji, Weizhen
; TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding Th
; FILE REFERENCE: 21402-141 CJP
; CURRENT APPLICATION NUMBER: US/10/096.625
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/972,211
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 1763
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (141)..(1715)
; US-10-096-625-29

Query Match          99.9%; Score 1576.4; DB 16; Length 1763;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGGCCACCTTGTCGCGACGACAGCTGTCCAAAGTGGCCACGCTCGGTGTCCAAACAGTCC 60
DB      141 ATGGCCACCTTGTCGCGACGACAGCTGTCCAAAGTGGCCACGCTCGGTGTCCAAACAGTCC 200

QY      61  CAGGCCAAGATGAGCGGCATGTTCCGACGATGGGTTTCAGCGCGGCCAGGATGAGGAG 120
DB      201 CAGGCCAAGATGAGCGGCATGTTCCGACGATGGGTTTCAGCGCGGCCAGGATGAGGAG 260

QY      121 GCGGTGGGCTTCGCGCATTCGACGACCTTCGACTTTGAGCACCGCCAGGGGCTTCGACATG 180
DB      261 GCGGTGGGCTTCGCGCATTCGACGACCTTCGACTTTGAGCACCGCCAGGGGCTTCGACATG 320

QY      181 GACATCTCTGAAGCCGAGGAGAGCCCTTCGCGGGAAGAGGCGCTGAGCGCCCTGCGAG 240
DB      321 GACATCTCTGAAGCCGAGGAGAGCCCTTCGCGGGAAGAGGCGCTGAGCGCCCTGCGAG 380

QY      241 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTGCGCGCCTCCGCGCTCCAAAGGAC 300

381  GGAGACATCCATTATCAGCGAGGCGGAGCTCTCTGCGCGCCTCCGCGCTCCAAAGGAC 440
301  CAGGTGGAGGTGGTGGCGGAATTCGGGGGCCACGACAAAGCCCAAAATACAGGGGTGGAG 360
441  CAGGTGGAGGTGGTGGCGGAATTCGGGGGCCACGACAAAGCCCAAAATACAGGGGTGGAG 500
361  GCAGGCTGGAACTGACCAACGCGCATCCAGGCGCATGTTCTGCTGGGCGCTACCTACGCC 420
501  GCAGGCTGGAACTGACCAACGCGCATCCAGGCGCATGTTCTGCTGGGCGCTACCTACGCC 560
421  ATCTGTGACGCGCGCTACCTGGGGTGTGTTTCTCATCATCTTTCGCGCGCTGTGTGTGCTG 480
561  ATCTGTGACGCGCGCTACCTGGGGTGTGTTTCTCATCATCTTTCGCGCGCTGTGTGTGCTG 620
481  TACACCGGCAAGATTCCTCATCGCGTGTGTAAGAGAGAAATGAAGACGCGAGGTGGTG 540
621  TACACCGGCAAGATTCCTCATCGCGTGTGTAAGAGAGAAATGAAGACGCGAGGTGGTG 680
541  CGGCTGGGAGCTCGTACGCGGCGCATAGCAACGCGCTGTGCGCGCGCGCTTCCCAACG 600
681  CGGCTGGGAGCTCGTACGCGGCGCATAGCAACGCGCTGTGCGCGCGCGCTTCCCAACG 740
601  CTGGCGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGAGCTGCATCCTG 660
741  CTGGCGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGAGCTGCATCCTG 800
661  TAGCTGTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGGTGCGCGGTGTGCGAG 720
801  TAGCTGTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGGTGCGCGGTGTGCGAG 860
721  AAGTCTGTGTCATTATCCGACGCGGTGCTGTGCTGCGCTTGGCGCTTCTTAAAGAACCTC 780
861  AAGTCTGTGTCATTATCCGACGCGGTGCTGTGCTGCGCTTGGCGCTTCTTAAAGAACCTC 920
781  AAGCGCGTGTCCAAAGTTCAGTCTGTGCTGTGCACCTCTGGCCACCTTCGTTCATCAATATC 840
921  AAGCGCGTGTCCAAAGTTCAGTCTGTGCTGTGCACCTCTGGCCACCTTCGTTCATCAATATC 980
841  GTCATAGCCTACTGTCTATCGCGGGCGCGCATGTCGGGCTGGGAGAGGTCAGTGTCTAC 900
981  GTCATAGCCTACTGTCTATCGCGGGCGCGCATGTCGGGCTGGGAGAGGTCAGTGTCTAC 1040
901  ATCGAGCTCAAGAAAGTTCCTCCATTCGATGGGATCATCGTGTTCAGCTACAGCTCTCAG 960
1041  ATCGAGCTCAAGAAAGTTCCTCCATTCGATGGGATCATCGTGTTCAGCTACAGCTCTCAG 1100
961  ATCTTCTGCTTCGCTGGAGGCAATATGCAGCAGCCAGCGAGTTCCTACTGCATGATG 1020
1101  ATCTTCTGCTTCGCTGGAGGCAATATGCAGCAGCCAGCGAGTTCCTACTGCATGATG 1160
1021  AACTGACGCACATCGCAGCTGCGTCTCAAGGGCTCTTTCGCGCTGCGCTACCTC 1080
1161  AACTGACGCACATCGCAGCTGCGTCTCAAGGGCTCTTTCGCGCTGCGCTACCTC 1220
1081  ACCTGGGCGCAGAGACCAAGAGGTCATACCGGATAACCTGCCCGGCTCCATCCGCGCC 1140
1221  ACCTGGGCGCAGAGACCAAGAGGTCATACCGGATAACCTGCCCGGCTCCATCCGCGCC 1280
1141  GTGGTCAAACATCTTCTGCTGGCCAAAGGCGCTGTTCTCTATCTCTGCGCATTTTGGCC 1200
1281  GTGGTCAAACATCTTCTGCTGGCCAAAGGCGCTGTTCTCTATCTCTGCGCATTTTGGCC 1340
1201  GCTGTGAGGTGCTGAGAAAGTGTCTTTCCAGGAAGGAGCGCGCGCTTTTTCCCGGCC 1260
1341  GCTGTGAGGTGCTGAGAAAGTGTCTTTCCAGGAAGGAGCGCGCGCTTTTTCCCGGCC 1400
1261  TGTACAGCGGCGAGCGGCGCTGAAGTCTTGGGGCTGACGCTGCGCTGCGCGCTGCTC 1320
1401  TGTACAGCGGCGAGCGGCGCTGAAGTCTTGGGGCTGACGCTGCGCTGCGCGCTGCTC 1460
1321  GTCTTCAAGCTGCTATGGCCATTTATGTGCGCGACTTTCGCGCTGCTCATGCGGCTCACC 1380
```

;; PRIOR FILING DATE: 2001-06-08
;; NUMBER OF SEQ ID NOS: 198
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 29
;; LENGTH: 1763
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (141)...(1715)
US-09-972-211-29

Query Match 99.9%; Score 1576.4; DB 11; Length 1763;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTGTCTGCGAGCAAGCTGTCCAAAGTGGCCACGTCGCTGTCGAAAGTCC 60
DB 141 ATGGCCACCTTGTCTGCGAGCAAGCTGTCCAAAGTGGCCACGTCGCTGTCGAAAGTCC 200

QY 61 CAGGCCAAGATGAGCGGCATGTTCCGACGATGGGTTTTCAGGCGGCCACGATGAGGAG 120
DB 201 CAGGCCAAGATGAGCGGCATGTTCCGACGATGGGTTTTCAGGCGGCCACGATGAGGAG 260

QY 121 GCGGTGGGCTTCGCGCATTCGACGACCTCGACTTTGACGACCGCGGCGCTGCGAGATG 180
DB 261 GCGGTGGGCTTCGCGCATTCGACGACCTCGACTTTGACGACCGCGGCGCTGCGAGATG 320

QY 181 GACATCCTGAAGCCGAGGAGAGCCCTCGGCGGACGAGGCGCTGGAAGCGCCGCTCGAG 240
DB 321 GACATCCTGAAGCCGAGGAGAGCCCTCGGCGGACGAGGCGCTGGAAGCGCCGCTCGAG 380

QY 241 GGAGACATCATATCAGCGAGCGGAGCGGAGCTCTCGCGCCCTCGGCTCCGAGGAC 300
DB 381 GGAGACATCATATCAGCGAGCGGAGCGGAGCTCTCGCGCCCTCGGCTCCGAGGAC 440

QY 301 CAGGTGGGAGTGGTGGCGAATTCGGGGGCCACGACCAAGCCCAAAATCAAGCGGCGGGAG 360
DB 441 CAGGTGGGAGTGGTGGCGAATTCGGGGGCCACGACCAAGCCCAAAATCAAGCGGCGGGAG 500

QY 361 GCAGGTGGAACGTGACCAACGCGCATCCAGGGCATGTTCTGCTGGGCGCTTACCTACGCC 420
DB 501 GCAGGTGGAACGTGACCAACGCGCATCCAGGGCATGTTCTGCTGGGCGCTTACCTACGCC 560

QY 421 ATCTCTCAGCGGCGTACCTGGGTTGTTCTCATCATCTTCGCGCGCTGTGTGCTGC 480
DB 561 ATCTCTCAGCGGCGTACCTGGGTTGTTCTCATCATCTTCGCGCGCTGTGTGCTGC 620

QY 481 TACACCGGCAAGATCCTCATTCGCGTGCCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 540
DB 621 TACACCGGCAAGATCCTCATTCGCGTGCCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 680

QY 541 CGCGTGGGACATCGTGGCCATAGCCCATAGCCCTGCTGCGCCCGCGCTTCCCAACG 600
DB 681 CGCGTGGGACATCGTGGCCATAGCCCATAGCCCTGCTGCGCCCGCGCTTCCCAACG 740

QY 601 CTGGGCGCGGAGTGTGAGCGATCATCGAGTGTGATGATGATGATGATGATGATGATG 660
DB 741 CTGGGCGCGGAGTGTGAGCGATCATCGAGTGTGATGATGATGATGATGATGATGATG 800

QY 661 TAGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
DB 801 TAGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 860

QY 721 AAGTCTGTGTCATTAATGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 861 AAGTCTGTGTCATTAATGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 920

QY 781 AAGGCGGTCGCAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 921 AAGGCGGTCGCAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 980

QY 841 GTCATAGCTTACTGTCTATTCGCGGGCGCGACTGGGCGCTGGGAGAGGTCAAGTTCCTAC 900

RESULT 6

US-10-096-625-29
; Sequence 29, Application US/10096625
; Publication No. US20040068095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burghess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grose, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R

DB 981 GTCATAGCTTACTGTCTATTCGCGGCGCGACTGGGCTGGGAGAGGTCAAGTTCCTAC 1040

QY 901 ATCGACGTCAAGAGTTCCCATCTCCATTGGCATCATCGTGTTCAGTACACGCTCTCAG 960

DB 1041 ATCGACGTCAAGAGTTCCCATCTCCATTGGCATCATCGTGTTCAGTACACGCTCTCAG 1100

QY 961 ATCTTCTGCTCTGCTGAGGGCAATATGACAGAGCCAGCGAGTTCACCTGATGATG 1020

DB 1101 ATCTTCTGCTCTGCTGAGGGCAATATGACAGAGCCAGCGAGTTCACCTGATGATG 1160

QY 1021 AACTGACGACATGACGACCTGCTCAAGGGCTCTTCGCGCTCGTGGCTACCTC 1080

DB 1161 AACTGACGACATGACGACCTGCTCAAGGGCTCTTCGCGCTCGTGGCTACCTC 1220

QY 1081 ACCTGGGCGGACGAGACCAAGAGGTCTACAGGATAACCTGCCGGGTCCATCCGCGCC 1140

DB 1221 ACCTGGGCGGACGAGACCAAGAGGTCTACAGGATAACCTGCCGGGTCCATCCGCGCC 1280

QY 1141 GTGGTCAACATCTTTCTGCTGCGCAAGCGCTGTTGCTCTATCTCTGCTCTTCTTGGC 1200

DB 1281 GTGGTCAACATCTTTCTGCTGCGCAAGCGCTGTTGCTCTATCTCTGCTCTTCTTGGC 1340

QY 1201 GCTGTGAGGTGCTGAGAGTCTCTTCAGAGAGGAGCGCGCTTTTTCGCGGCT 1260

DB 1341 GCTGTGAGGTGCTGAGAGTCTCTTCAGAGAGGAGCGCGCTTTTTCGCGGCT 1400

QY 1261 TGCTACAGCGGCGAGCGGCGCTGAGGTCCTGGGGGCTGAGCGCTGCGCGCTCGTC 1320

DB 1401 TGCTACAGCGGCGAGCGGCGCTGAGGTCCTGGGGGCTGAGCGCTGCGCGCTCGTC 1460

QY 1321 GTCTTACGCTGCTCATGGCCATTTATGTGCGGCACTTCGCGCTGCTCATGGGCTCAC 1380

DB 1461 GTCTTACGCTGCTCATGGCCATTTATGTGCGGCACTTCGCGCTGCTCATGGGCTCAC 1520

QY 1381 GCGACCTTACGCGGCGCGGCTCTGTTTCTGCTGCGGAGCTCTTTCACCTGCGGCTG 1440

DB 1521 GCGACCTTACGCGGCGCGGCTCTGTTTCTGCTGCGGAGCTCTTTCACCTGCGGCTG 1580

QY 1441 CTCTGGCGACAGCTGCTGCGGCGGCTCTTTCGAGCTGCGCATCTTCGTCATCGGC 1500

DB 1581 CTCTGGCGACAGCTGCTGCGGCGGCTCTTTCGAGCTGCGCATCTTCGTCATCGGC 1640

QY 1501 GGCATCTGAGGCTGCGGCTTCTGCACTCCCTCGAGGGCTCATCGAAGCTACCGA 1560

DB 1641 GGCATCTGAGGCTGCGGCTTCTGCACTCCCTCGAGGGCTCATCGAAGCTACCGA 1700

QY 1561 ACCAAGCGGAGGACTAG 1578

DB 1701 ACCAAGCGGAGGACTAG 1718

[illegible]

```

RESULT 5
US-09-972-211-29
: Sequence 29, Application US/09972211
: Publication No. US20040048245A1
: GENERAL INFORMATION:
: APPLICANT: Shinkets, Richard A
: APPLICANT: Taupier Jr, Raymond J
: APPLICANT: Burgess, Catherine E
: APPLICANT: Zerhusen, Bryan D
: APPLICANT: Mezes, Peter S
: APPLICANT: Rastelli, Luca
: APPLICANT: Malyankar, Uriel M
: APPLICANT: Grosse, William M
: APPLICANT: Alsobrook II, John P
: APPLICANT: Lepley, Denise M
: APPLICANT: Spytek, Kimberly Ann
: APPLICANT: Li, Li
: APPLICANT: Edinger, Shlomit
: APPLICANT: Gerlach, Valerie
: APPLICANT: Ellerman, Karen
: APPLICANT: MacDougall, John R
: APPLICANT: Gunther, Erik
: APPLICANT: Millet, Isabelle
: APPLICANT: Stone, David J
: APPLICANT: Smithson, Glenda
: APPLICANT: Szekeres Jr, Edward S
: TITLE OF INVENTION: No. US20040048245A1el Humana
: FILE REFERENCE: 21402-141
: CURRENT APPLICATION NUMBER: US/09/972,211
: CURRENT FILING DATE: 2001-10-05
: PRIOR APPLICATION NUMBER: 60/238,325
: PRIOR FILING DATE: 2000-10-05
: PRIOR APPLICATION NUMBER: 60/238,323
: PRIOR FILING DATE: 2000-10-05
: PRIOR APPLICATION NUMBER: 60/238,400
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/238,397
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/238,401
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/238,379
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/238,402
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 30/238,384
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/238,373
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/238,372
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/238,383
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/238,382
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/275,892
: PRIOR FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: 60/296,860

```

QY 421 ATCTGCGAGCGGCTACTCGGGTGTGTTCTCATCATCTTCGCGCGGTTGTGTGCTGC 480
Db 829 ATCTGCGAGCGGCGGTACTCGGGTGTGTTCTCATCATCTTCGCGCGGTTGTGTGCTGC 888
QY 481 TACACCGGCAAGATCCTCATCGCTGCTGTACGAGGAGATGAAGACGGGAGGTGGTG 540
Db 889 TACACCGGCAAGATCCTCATCGCTGCTGTACGAGGAGATGAAGACGGGAGGTGGTG 948
QY 541 CCGCTGCGGAGCTGTAGTGCCCATAGCCAAAGCGCTGTGCGCCCGCGGCTTCCCAACG 600
Db 949 CCGCTGCGGAGCTGTAGTGCCCATAGCCAAAGCGCTGTGCGCCCGCGGCTTCCCAACG 1008
QY 601 CTGGGCGGCGAGTGTGAAGCTAGCGAGATCATCGAGCTGTGATGAGTGATCTCTG 660
Db 1009 CTGGGCGGCGAGTGTGAAGCTAGCGAGATCATCGAGCTGTGATGAGTGATCTCTG 1068
QY 661 TACGTGGTGGTGAAGTGAAGCTCATGATCAACAGCTTCCGGGCTGCGGTGTCGAG 720
Db 1069 TACGTGGTGGTGAAGTGAAGCTCATGATCAACAGCTTCCGGGCTGCGGTGTCGAG 1128
QY 721 AAGTCTGTGTCATATGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 1129 AAGTCTGTGTCATATGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1188
QY 781 AAGGCGGCTGTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 1189 AAGGCGGCTGTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
QY 841 GTATAGGCTACTGTCTATCGGGGCGGCGGCTGCGGCTGCGGAGAGGTCAAGTCTTAC 900
Db 1249 GTATAGGCTACTGTCTATCGGGGCGGCGGCTGCGGCTGCGGAGAGGTCAAGTCTTAC 1308
QY 901 ATGAGCTCAAGAGTTCCTCATATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 1309 ATGAGCTCAAGAGTTCCTCATATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1368
QY 961 ATCTCTCGCTGCTGCTGAGGCGCATATGCGAGCGGCGGAGTTCACCTGCTGATG 1020
Db 1369 ATCTCTCGCTGCTGCTGAGGCGCATATGCGAGCGGCGGAGTTCACCTGCTGATG 1428
QY 1021 AACTGAGCGACATCGAGCGCTGCTGCTCAAGGGCTCTTCGCGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1429 AACTGAGCGACATCGAGCGCTGCTGCTCAAGGGCTCTTCGCGCTGCTGCTGCTGCTGCTGCTG 1488
QY 1081 ACTTGGGCGACGAGACCAAGAGGTTCATCGGATTAACCTGCGCGGCTCCATCGCGGC 1140
Db 1489 ACTTGGGCGGACGAGACCAAGAGGTTCATCGGATTAACCTGCGCGGCTCCATCGCGGC 1548
QY 1141 GTGGTCAACATCTTCTGCTGCGCAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 1549 GTGGTCAACATCTTCTGCTGCGCAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1608
QY 1201 GCTGTCAGGCTGCTGAGAGTGCCTCTTCCAGGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db 1609 GCTGTCAGGCTGCTGAGAGTGCCTCTTCCAGGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1668
QY 1261 TGCTACAGGCGGCGGCGGCTGAGAGTCTGCGGCGCTGAGCGTGGGCTGCGGCTGCTGCTGCTGCTGCTG 1320
Db 1669 TGCTACAGGCGGCGGCGGCTGAGAGTCTGCGGCGCTGAGCGTGGGCTGCGGCTGCTGCTGCTGCTGCTG 1728
QY 1321 GTCTTCAAGCTGCTCATGCGCATTTATGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTG 1380
Db 1729 GTCTTCAAGCTGCTCATGCGCATTTATGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTG 1788
QY 1381 GGAGGCTTCAAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
Db 1789 GGAGGCTTCAAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1848
QY 1441 CTCTGCGCAAGCTGTGTGGCAACCAAGTCTTTCGAGTTCGCGCATCTTTCGCTGCTGCTGCTGCTGCTGCTG 1500
Db 1849 CTCTGCGCAAGCTGTGTGGCAACCAAGTCTTTCGAGTTCGCGCATCTTTCGCTGCTGCTGCTGCTGCTGCTG 1908

QY 1501 GGCACTCTGACGCTGTCGGCTTCGTCACCTCCTCGAGGCGCTCATCGAAGCCTACCGA 1560
Db 1909 GGCACTCTGACGCTGTCGGCTTCGTCACCTCCTCGAGGCGCTCATCGAAGCCTACCGA 1968
QY 1561 ACCAACGGGAGGACTAG 1578
Db 1969 ACCAACGGGAGGACTAG 1986
RESULT 4
US-10-789-241-9
; Sequence 9, Application US/10789241
; Publication No. US2004018032A1
; GENERAL INFORMATION:
; APPLICANT: Powell, Douglas M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: AIDS AND HIV-RELATED DISORDERS USING 9145, 1725, 311, 837,
; TITLE OF INVENTION: 58305, 156, 14175, 50352, 32678, 5560, 7240, 8855, 12396,
; TITLE OF INVENTION: 12397, 13644, 19938, 2077, 1735, 1786, 10220, 17822, 33945,
; TITLE OF INVENTION: 43748, 47161, 81982 OR 46777
; FILE REFERENCE: MPI03-041P1RNMNIM
; CURRENT APPLICATION NUMBER: US/10/789,241
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/454,202
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/456,326
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/465,240
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 60/475,233
; PRIOR FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 60/478,952
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/487,836
; PRIOR FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/500,111
; PRIOR FILING DATE: 2003-09-04
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2585
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (248)...(1825)
US-10-789-241-9
Query Match 100.0%; Score 1578; DB 17; Length 2585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCACCTTCTCGGAGCAAGCTGTCCAACTGGCCACGTCCTGTCCTGTCCTCAACAGTCC 60
Db 248 ATGGCCACCTTCTCGGAGCAAGCTGTCCAACTGGCCACGTCCTGTCCTGTCCTCAACAGTCC 307
QY 61 CAGGCCAAGATCAGCGGCTGTTCCAGGATGGTTCAGCGCGCCACCGATGAGGAG 120
Db 308 CAGGCCAAGATCAGCGGCTGTTCCAGGATGGTTCAGCGCGCCACCGATGAGGAG 367
QY 121 GCGGTGGGCTTCGCGCATTTGCGACGACCTTCGATTTGAGCACCGCCAGGCGCTGCAATG 180
Db 368 GCGGTGGGCTTCGCGCATTTGCGACGACCTTCGATTTGAGCACCGCCAGGCGCTGCAATG 427
QY 181 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGAGAGGCGCTGGAAGCGCCCTCGAG 240
Db 428 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGAGAGGCGCTGGAAGCGCCCTCGAG 487
QY 241 GGAGACATCTTATTCAGCGAGGAGCGAGTCTCTGCGCGCTTCCGGCTCCAAAGGAC 300
Db 489 GGAGACATCTTATTCAGCGAGGAGCGAGTCTCTGCGCGCTTCCGGCTCCAAAGGAC 547
QY 301 CAGGTGGGAGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCATCGGCGTGGAG 360

Db 254 GAGACATCCATTATACAGGAGGAGCGAGCTCTCTCGCGCCCTCCGGCTCCAGGAC 313
QY 301 CAGGTGGAGGTGGTGGCAATTCGGGGGCCACGACAGGCCAAATACAGGCGTGGGAG 360
Db 314 CAGGTGGAGGTGGTGGCAATTCGGGGGCCACGACAGGCCAAATACAGGCGTGGGAG 373
QY 361 GCAGGCTGGAACGTCGACCAACGCCATCCAGGCGATGTTCTGTGTGGGSCCTACCCCTAGCC 420
Db 374 GCAGGCTGGAACGTCGACCAACGCCATCCAGGCGATGTTCTGTGTGGGSCCTACCCCTAGCC 433
QY 421 ATCTGACCGCGCGCTACCTCGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGTGCTGC 480
Db 434 ATCTGACCGCGCGCTACCTCGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGTGCTGC 493
QY 481 TACACCGGCAAGATCTCATCTGCGGTGCTGTACGAGGAGAAATGAAGCGCGAGGTGGT 540
Db 494 TACACCGGCAAGATCTCATCTGCGGTGCTGTACGAGGAGAAATGAAGCGCGAGGTGGT 553
QY 541 CGGTGCGGACTCGTACGTGGCCATAGCAACGCCCTGCTGCGCGCCCGCTTCCCAACG 600
Db 554 CGGTGCGGACTCGTACGTGGCCATAGCAACGCCCTGCTGCGCGCCCGCTTCCCAACG 613
QY 601 CTGGCGCGCGAGTGGTGAACGTCAGGCGAGATCATCGAGCTGGTGTGATGAGCTGCATCCTG 660
Db 614 CTGGCGCGCGAGTGGTGAACGTCAGGCGAGATCATCGAGCTGGTGTGATGAGCTGCATCCTG 673
QY 661 TAGTGGTGTGAGTGGCAACCTCATGTACACAGCTTCCGGGGTGCCTGTGTCGAG 720
Db 674 TAGTGGTGTGAGTGGCAACCTCATGTACACAGCTTCCGGGGTGCCTGTGTCGAG 733
QY 721 AAGTCTGTGTCATTAATGCGCACGGCGTCTGCTGCTTGGCGCTTCCCTTAAGAACCTC 780
Db 734 AAGTCTGTGTCATTAATGCGCACGGCGTCTGCTGCTTGGCGCTTCCCTTAAGAACCTC 793
QY 781 AAGGCCGTGTCGAAGTTCAAGTCTGCTGTGCACCTCTGCGCCACCTTCGTCAATATCCTG 840
Db 794 AAGGCCGTGTCGAAGTTCAAGTCTGCTGTGCACCTCTGCGCCACCTTCGTCAATATCCTG 853
QY 841 GTCATAGCTACTGTCTATCGGGGCGGCGACTGGGCTGGAGAGGTCAAGTTCATC 900
Db 854 GTCATAGCTACTGTCTATCGGGGCGGCGACTGGGCTGGAGAGGTCAAGTTCATC 913
QY 901 ATCGAGCTCAAGAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGCTCTCAG 960
Db 914 ATCGAGCTCAAGAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGCTCTCAG 973
QY 961 ATCTTCTGCTTTCGTGGAGGGCAATATGACAGCGCCAGCGAGTTCCACTGCATGATG 1020
Db 974 ATCTTCTGCTTTCGTGGAGGGCAATATGACAGCGCCAGCGAGTTCCACTGCATGATG 1033
QY 1021 AACTGAGCGACATCGAGCTCGTGTCAAGGGCTCTTTCGCGCTCGTGTGCTACCTC 1080
Db 1034 AACTGAGCGACATCGAGCTCGTGTCAAGGGCTCTTTCGCGCTCGTGTGCTACCTC 1093
QY 1081 ACCTGGGCGACAGAGAGGTTCATCAGCGATAA CTTGCGCGGCTCCATCCGCGC 1140
Db 1094 ACCTGGGCGACAGAGAGGTTCATCAGCGATAA CTTGCGCGGCTCCATCCGCGC 1153
QY 1141 GTGGTCAACATCTTTCTGTGGCAAGGGCTGTGTCTTATCTCTGCGCATTTCTTTGGC 1200
Db 1154 GTGGTCAACATCTTTCTGTGGCAAGGGCTGTGTCTTATCTCTGCGCATTTCTTTGGC 1213
QY 1201 GCTGTGAGGTGTGAGAGTTCGCTCTTCAGGAAAGCGAGCGCGCTTTTTCGCGGC 1260
Db 1214 GCTGTGAGGTGTGAGAGTTCGCTCTTCAGGAAAGCGAGCGCGCTTTTTCGCGGC 1273
QY 1261 TGCTACAGCGGACAGCGCGCTGAGTCTGTGGGGCTGACGCTGCGCTGCGGCTCCTC 1320
Db 1274 TGCTACAGCGGACAGCGCGCTGAGTCTGTGGGGCTGACGCTGCGCTGCGGCTCCTC 1333
QY 1321 GTCTTACGCTGCTCATGGCCATTTATGTGCCCACTTTCGCGCTGCTCATGGGCTTCACC 1380

Db 1334 GTCTTACGCTGCTCATGGCCATTTATGTGCGCACTTTCGCGCTGCTCATGGGCTCAC 1393
QY 1381 GGAGAGCTCAAGGCGCGGCTCTGTCTTCTTGTCTGCCAGGCTCTTTCACCTGCGGCTG 1440
Db 1394 GGAGAGCTCAAGGCGCGGCTCTGTCTTCTTGTCTGCCAGGCTCTTTCACCTGCGGCTG 1453
QY 1441 CTCTGCGCAAGCTGTGTGGCACCAAGTCTTCTTCGAGCTGCGCATCTTCGTTCATCGGC 1500
Db 1454 CTCTGCGCAAGCTGTGTGGCACCAAGTCTTCTTCGAGCTGCGCATCTTCGTTCATCGGC 1513
QY 1501 GGCACTGCAAGCTGTTCGCGCTTTCGTGCACTCTCCCTCGAGGGCTCATCGAAGCTTACCGA 1560
Db 1514 GGCACTGCAAGCTGTTCGCGCTTTCGTGCACTCTCCCTCGAGGGCTCATCGAAGCTTACCGA 1573
QY 1561 ACCAACCGGAGGACTAG 1578
Db 1574 ACCAACCGGAGGACTAG 1591

RESULT 3
US-09-940-919-3
; Sequence 3, Application US/09940919
; Patent No. US20020082390A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: No. US20020082390A1e1 Human GABA Transporter Protein and Polynuc
; FILE REFERENCE: LEX-0228-USA
; CURRENT APPLICATION NUMBER: US/09/940,919
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/230,178
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2261
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-940-919-3

Query Match 100.0%; Score 1578; DB 9; Length 2261;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTGTCTCGCAGCAAGCTGTCCAACTGGCGCACGTCCTGTCTCCAACTGTC 60
Db 409 ATGGCCACCTTGTCTCGCAGCAAGCTGTCCAACTGGCGCACGTCCTGTCTCCAACTGTC 468
QY 61 CAGGCCAAGATGAGCGGCATGTTGCGCAGGATGGGTTTTTCAGCGCGCCACGAGTAGGAG 120
Db 469 CAGGCCAAGATGAGCGGCATGTTGCGCAGGATGGGTTTTTCAGCGCGCCACGAGTAGGAG 528
QY 121 GCGGTGGGCTTCGCGCATTTGCGACACCTTCGACCTTCGACCTTCGACCTTCGACCTTCGACATG 180
Db 529 GCGGTGGGCTTCGCGCATTTGCGACACCTTCGACCTTCGACCTTCGACCTTCGACATG 588
QY 181 GACATCTCTGAAACCGGAGGAGAGCCCTGCGGGGACGAGGGGCTGAAAGCGCCGCTCGAG 240
Db 589 GACATCTCTGAAACCGGAGGAGAGCCCTGCGGGGACGAGGGGCTGAAAGCGCCGCTCGAG 648
QY 241 GGAGACATCCATTATCAGCGAGGACGAGGAGTCTCTTCGCGGCTTCGCGCTTCGAGGAC 300
Db 649 GGAGACATCCATTATCAGCGAGGACGAGGAGTCTCTTCGCGGCTTCGCGCTTCGAGGAC 708
QY 301 CAGGTGGGAGGTGGCGAAATTCGGGGCGCACGACCAAGCCCAAATACAGCGCTGGGAG 360
Db 709 CAGGTGGGAGGTGGCGAAATTCGGGGCGCACGACCAAGCCCAAATACAGCGCTGGGAG 768
QY 361 CGAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTTGTGTGCTGGGCTTACCTTACGCC 420
Db 769 CGAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTTGTGTGCTGGGCTTACCTTACGCC 828

```
QY 121 GCGGTGGGCTTCGCGATTCGACGACCTCGACTTTGAGCACCGCGCGGCTTCAGATG 180
Db 121 GCGGTGGGCTTCGCGATTCGACGACCTCGACTTTGAGCACCGCGCGGCTTCAGATG 180
QY 181 GACATCTGTAAGAGCGAGGAGAGCCCTGCGGGGACAGAGGGCGCTGAAGCGCCGCTCGAG 240
Db 181 GACATCTGTAAGAGCGAGGAGAGCCCTGCGGGGACAGAGGGCGCTGAAGCGCCGCTCGAG 240
QY 241 GGAGACATCCATTATCAGCGAGGAGCGAGGCTCTCTGCGCGCCCTCGGGCTCCAAGGAC 300
Db 241 GGAGACATCCATTATCAGCGAGGAGCGAGGCTCTCTGCGCGCCCTCGGGCTCCAAGGAC 300
QY 301 CAGGTGGGAGGTGGTGGGGAATTCGGGGGCGCACGACAAGCGCCCAAAATCACGCGCGTGGGAG 360
Db 301 CAGGTGGGAGGTGGTGGGGAATTCGGGGGCGCACGACAAGCGCCCAAAATCACGCGCGTGGGAG 360
QY 361 GCAGGCTGGAAGCGTGAACCAACCCATCCAGGGCATGTTGCTGCTGGGCTTACCCTACGCC 420
Db 361 GCAGGCTGGAAGCGTGAACCAACCCATCCAGGGCATGTTGCTGCTGGGCTTACCCTACGCC 420
QY 421 ATCTGTCACGGCGGTACTCTGCGGTGTTGTTCTCATCATCTTCGCGCGCTGTTGTCGTCG 480
Db 421 ATCTGTCACGGCGGTACTCTGCGGTGTTGTTCTCATCATCTTCGCGCGCTGTTGTCGTCG 480
QY 481 TACACCGGCAAGATCTCATCGCTGCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 540
Db 481 TACACCGGCAAGATCTCATCGCTGCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 540
QY 541 CCGGTGCGGGAATCTGAGTGGCCATAGCCAAACGCTGTCGCGCCCGCGCTTCCCAACG 600
Db 541 CCGGTGCGGGAATCTGAGTGGCCATAGCCAAACGCTGTCGCGCCCGCGCTTCCCAACG 600
QY 601 CTGGGGGCGGAGTGGTGAACCTAGCGAGATCATCGAGCTGGTGCATGACGTGCATCTG 660
Db 601 CTGGGGGCGGAGTGGTGAACCTAGCGAGATCATCGAGCTGGTGCATGACGTGCATCTG 660
QY 661 TACGTGCTGCTGAGTGGCAACCTCATGTACAAACAGCTTCCCGGGGCTGCGGTGTCGAG 720
Db 661 TACGTGCTGCTGAGTGGCAACCTCATGTACAAACAGCTTCCCGGGGCTGCGGTGTCGAG 720
QY 721 AAGTCTGCTCATTAATGCGCAAGCGCGTGTGCTGCTGCGCTTCCCTTGAAGAACCTC 780
Db 721 AAGTCTGCTCATTAATGCGCAAGCGCGTGTGCTGCTGCGCTTCCCTTGAAGAACCTC 780
QY 781 AAGGCGGTGTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 840
Db 781 AAGGCGGTGTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 840
QY 841 GTCATAGCCTACTGTCTATCGGGGCGCGGAGTGGGCTGGGAGAGGTCAAGTTCATC 900
Db 841 GTCATAGCCTACTGTCTATCGGGGCGCGGAGTGGGCTGGGAGAGGTCAAGTTCATC 900
QY 901 ATCGAGCTCAAGAAGTTCCTCATTCATTTGAGCATCATCGTGTTCAGCTTACAGTTCAG 960
Db 901 ATCGAGCTCAAGAAGTTCCTCATTCATTTGAGCATCATCGTGTTCAGCTTACAGTTCAG 960
QY 961 ATCTTCTGCTGCTGCTGAGGGGCAATATGCAAGAGGAGTTCAGTTCAGTTCAGTTCAG 1020
Db 961 ATCTTCTGCTGCTGCTGAGGGGCAATATGCAAGAGGAGTTCAGTTCAGTTCAGTTCAG 1020
QY 1021 AACTGAGCGCATCATCGAGCTGCGTGTCTCAAGGGGCTCTTCGCGCTGCTGCGCTTACCTC 1080
Db 1021 AACTGAGCGCATCATCGAGCTGCGTGTCTCAAGGGGCTCTTCGCGCTGCTGCGCTTACCTC 1080
QY 1081 ACTTGGGCGGACAGACCAAGAGGTTCATCAAGGATAAATCTGCGGGCTTCATTCGCGGCTC 1140
Db 1081 ACTTGGGCGGACAGACCAAGAGGTTCATCAAGGATAAATCTGCGGGCTTCATTCGCGGCTC 1140
QY 1141 GTGGTCAACATCTTCTGCTGCGCAAGGCGCTGTTGCTATCTCTGCTGCTGCTTCTTGGC 1200
Db 1141 GTGGTCAACATCTTCTGCTGCGCAAGGCGCTGTTGCTATCTCTGCTGCTGCTTCTTGGC 1200
QY 1201 GCTGTCGAGGTGCTGAGAGTTCGCTTTCAGGAGGAGCGGCGCTTCTGCGCGCTTCTGCGG 1260
```

```
Db 1201 GCTGTCGAGGTGCTGAGAGTTCGCTTCTCAGGAAGCGAGCGCGCTTCTTCCCGGCC 1260
QY 1261 TGCTACAGCGGAGCGGCGCTGAACTCTTGGGGGCTGAGCTGCGGTGCGGTGCGCTGTC 1320
Db 1261 TGCTACAGCGGAGCGGCGCTGAACTCTTGGGGGCTGAGCTGCGGTGCGGTGCGCTGTC 1320
QY 1321 GTCTTACAGCTGCTCATGGCCATTTATGTCGCGCACTTCGCGCTGCTCATGGGCTCAC 1380
Db 1321 GTCTTACAGCTGCTCATGGCCATTTATGTCGCGCACTTCGCGCTGCTCATGGGCTCAC 1380
QY 1381 GGCAGCTTACAGCGGCGGCGCTCTGTTTCTGTCGCGGCTTCTTCACTGCGGCTG 1440
Db 1381 GGCAGCTTACAGCGGCGGCGCTCTGTTTCTGTCGCGGCTTCTTCACTGCGGCTG 1440
QY 1441 CTCTGCGGCAAGCTGCTGTCGCGCAAGTCTTCTTTCGAGCTGCGCATCTTCTGTCATCG 1500
Db 1441 CTCTGCGGCAAGCTGCTGTCGCGCAAGTCTTCTTTCGAGCTGCGCATCTTCTGTCATCG 1500
QY 1501 GGCATCTGCGGCTGTCGCGCTTCTGTCGCACTTCCCTCGAGGGCTTATCGAAGCTTACCG 1560
Db 1501 GGCATCTGCGGCTGTCGCGCTTCTGTCGCACTTCCCTCGAGGGCTTATCGAAGCTTACCG 1560
QY 1561 ACCAAGCGGAGGACTAG 1578
Db 1561 ACCAAGCGGAGGACTAG 1578

RESULT 2
US-09-871-503-1
; Sequence 1, Application US/09871503
; Patent No. US20020076758A1
; GENERAL INFORMATION:
; APPLICANT: George Christian Teretappen
; APPLICANT: Cinzia Felicita Sala
; TITLE OF INVENTION: POLYPEPTIDE
; FILE REFERENCE: QG 1013
; CURRENT APPLICATION NUMBER: US/09/871,503
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: GB 0013239.9
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1595
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14)..(1588)
US-09-871-503-1

Query Match 100.0%; Score 1578; DB 9; Length 1595;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCACTTGTCTCCGAGCAAGCTGTCCAAAGTGGGCAAGCTGCGTGTCCAAAGTTC 60
Db 14 ATGGCCACTTGTCTCCGAGCAAGCTGTCCAAAGTGGGCAAGCTGCGTGTCCAAAGTTC 73
QY 61 CAGGCCAAGATGAGCGGCTGTCGCCAGGATGGGTTTTTTCAGGGGCGCACCGATGAGGAG 120
Db 74 CAGGCCAAGATGAGCGGCTGTCGCCAGGATGGGTTTTTTCAGGGGCGCACCGATGAGGAG 133
QY 121 GCGGTGGGCTTTCGCGCATTTGCGACCTTCAGCTTTGAGCACCGCGGCTTGCAGATG 180
Db 134 GCGGTGGGCTTTCGCGCATTTGCGACCTTCAGCTTTGAGCACCGCGGCTTGCAGATG 193
QY 181 GACATCTGAAAGCGGAGGAGAGCCCTGCGGGGACGAGGGCGCTGAAAGCGCCCTCGAG 240
Db 194 GACATCTGAAAGCGGAGGAGAGCCCTGCGGGGACGAGGGCGCTGAAAGCGCCCTCGAG 253
QY 241 GGAGACATCCATTATCAGCGAGGAGCGGAGTTCCTCTGCGCGCTTCTGCGGCTTCAAGGAC 300
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 19:03:12 ; Search time 782 Seconds
(without alignments)
10347.002 Million cell updates/sec

Title: US-09-940-919-1

Perfect score: 1578

Sequence: 1 atggccacctgtctccgag.....gaaccaacggaggactag 1578

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1578	100.0	1578	9	US-09-940-919-1
2	1578	100.0	1595	9	US-09-871-503-1
3	1578	100.0	2261	9	US-09-940-919-3
4	1578	100.0	2585	17	US-10-789-241-9
5	1576.4	99.9	1763	11	US-09-972-211-29
6	1576.4	99.9	1763	16	US-10-096-625-29
7	1576.4	99.9	1800	14	US-10-239-566-1
8	1574.8	99.8	1734	11	US-09-972-211-31
9	1574.8	99.8	1734	11	US-09-972-211-35
10	1574.8	99.8	1734	11	US-09-972-211-37
11	1574.8	99.8	1734	11	US-09-972-211-39
12	1574.8	99.8	1734	16	US-10-096-625-31

13	1574.8	99.8	1734	16	US-10-096-625-35	Sequence 35, Appl
14	1574.8	99.8	1734	16	US-10-096-625-37	Sequence 37, Appl
15	1574.8	99.8	1734	16	US-10-096-625-39	Sequence 39, Appl
16	1567.4	99.3	1733	11	US-09-972-211-33	Sequence 33, Appl
17	1567.4	99.3	1733	16	US-10-096-625-33	Sequence 33, Appl
18	237	15.0	255	10	US-09-918-995-2561	Sequence 2561, Ap
19	177.2	11.2	498	10	US-09-918-995-26051	Sequence 26051, A
20	105	6.7	1746	16	US-10-425-114-2892	Sequence 2892, Ap
21	80.4	5.1	1045	17	US-10-767-701-14757	Sequence 14757, A
22	78.2	5.0	1819	17	US-10-437-963-896	Sequence 896, App
23	77.2	4.9	467	16	US-10-425-114-5154	Sequence 5154, Ap
24	76.4	4.8	1756	16	US-10-425-114-15506	Sequence 15506, A
25	76.4	4.8	1793	16	US-10-425-114-15832	Sequence 15832, A
26	75.6	4.8	692	17	US-10-767-701-9675	Sequence 9675, Ap
27	74.8	4.7	2100	17	US-10-437-963-330	Sequence 330, App
28	72	4.6	1289	17	US-10-437-963-97181	Sequence 97181, A
29	71	4.5	720	13	US-10-101-487-74	Sequence 74, Appl
30	71	4.5	720	13	US-10-101-487-76	Sequence 76, Appl
31	70.6	4.5	1452	17	US-10-437-963-69156	Sequence 69156, A
32	68.8	4.4	815	17	US-10-767-701-8980	Sequence 8980, Ap
33	68.8	4.4	1781	16	US-10-425-114-34120	Sequence 34120, A
34	67	4.2	967	17	US-10-437-963-97185	Sequence 97185, A
35	65	4.1	745	17	US-10-437-963-43816	Sequence 43816, A
36	65	4.1	1821	15	US-10-355-430-3	Sequence 3, Appl1
37	64.6	4.1	1212	16	US-10-260-238-267	Sequence 267, App
38	64.6	4.1	3264	17	US-10-437-963-43155	Sequence 43155, A
39	61	3.9	1626	17	US-10-437-963-32293	Sequence 32293, A
40	59.8	3.8	579	11	US-09-758-759-132	Sequence 132, App
41	59.8	3.8	109519	11	US-09-758-759-1	Sequence 1, Appl1
42	59.2	3.8	1767	17	US-10-437-963-92518	Sequence 92518, A
43	58	3.7	1407	17	US-10-437-963-94722	Sequence 94722, A
44	58	3.7	3489	11	US-09-894-273-1	Sequence 1, Appl1
45	58	3.7	3489	15	US-10-294-804-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1

US-09-940-919-1
; Sequence 1, Application US/09940919
; Patent No. US30020082390A1
; GENERAL INFORMATION:
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0228-USA
; CURRENT APPLICATION NUMBER: US/09/940,919
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/230,178
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-940-919-1

Query Match 100.0%; Score 1578; DB 9; Length 1578;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCCACCTTCTCGCAGCAAGTGTCCAACGTGGCCACGTCCGTGTCCAAACAGTCC	60
Db	1	ATGGCCACCTTCTCGCAGCAAGTGTCCAACGTGGCCACGTCCGTGTCCAAACAGTCC	60
QY	61	CAGGCCAAGATGAGCGGCATGTTCCCGCAGGATGGGTTTTCAGCGCGGCCACCGATGAGGAG	120
Db	61	CAGGCCAAGATGAGCGGCATGTTCCCGCAGGATGGGTTTTCAGCGCGGCCACCGATGAGGAG	120


```
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 20
; LENGTH: 32207
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20

Query Match      3.7%; Score 58; DB 3; Length 32207;
Best Local Similarity 52.5%; Pred. No. 0.0011;
Matches 127; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 1317 CGTCGCTTCACGCTGCTCATGGCCATTTATGTCGGCAGCTTCGGCGCTGCTCATGGGCT 1376
DB 20703 CATGGCCCCGTCCTCCTCGTCCTTCTTGTCCTCCCTCGTCATCCTCCTCGTCT 20762

QY 1377 CACGGGACGCTCACGGGCGCGGCTCTGTTCTTGCTGCCAGCCTCTTTACGCTGCG 1436
DB 20763 CATGTCCTCATCATGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20822

QY 1437 CTGCTCTGGCGCAAGCTGCTGTGGCAACCAAGTCTTTCGACGTCGCCATCTTGTGCTAT 1496
DB 20823 CCTCCTCGTCCCTCCTCTCGTCATCTCTCTCGTCATCTCTCTCTCTCTCTCTCTCT 20882

QY 1497 CGGCGGCATCTGACGGTGTCCGGCTTCTGTCACCTCCCTCGAGGGCTCATCGAAGCCTA 1556
DB 20883 CCTCCTCGTCATCTCTCTCGTCATCTCTCTCGTCATCTCTCTCTCTCTCTCTCTCT 20942

QY 1557 CC 1558
DB 20943 CC 20944
```

Search completed: November 3, 2004, 22:25:45
Job time : 166 secs

Db 1054 CC 1053

RESULT 13

US-08-770-379-20

; Sequence 20, Application US/08770379

; Patent No. 5849564

; GENERAL INFORMATION:

; APPLICANT: Chang, Yuan

; APPLICANT: Bohenzky, Roy A.

; APPLICANT: Russo, James J.

; APPLICANT: Edelman, Isidore S.

; APPLICANT: Moore, Patrick S.

; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED

; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/770,379

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 52342

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3207 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-770-379-20

Query Match 3.7%; Score 58; DB 2; Length 3207;

Best Local Similarity 52.5%; Pred. No. 0.0011;

Matches 127; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

Qy 1317 CGTCGCTTTCACGCTGCTCATGGCCATTATGCGCGCACTTCGCGCTGCTCATGGGCGCT 1376

Db 20703 CATGCCCGCGTCTCTCTGCT 20762

Qy 1377 CACGGCAGCTCACGGGCGCGGCGCTCTGTTTCTTGTCGCCAGGCTCTTTTACCTGGC 1436

Db 20763 CATGTCTCTCATCATGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20822

Qy 1437 CTTGCTCTGGCGCAAGTGTGTGGACCAAGTCTTCTTGAGGTGCGCATCTTGTCTCAT 1496

Db 20823 CTTCTCTGCT 20882

Qy 1497 CGGCGGCACTGTCAGGCGTGTCCGGCTTCGTGCACTCCCTTCGAGGGCTCATCGAAGCCTA 1556

Db 20883 CTTCTCTGCT 20942

Qy 1557 CC 1558

Db 20943 CC 20944

RESULT 14

US-08-757-669A-20

; Sequence 20, Application US/08757669A

; Patent No. 6183751

; GENERAL INFORMATION:

; APPLICANT: Chang, Yuan

; APPLICANT: Bohenzky, Roy A.

; APPLICANT: Russo, James J.

; APPLICANT: Edelman, Isidore S.

; APPLICANT: Moore, Patrick S.

; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS

; TITLE OF INVENTION: SEQUENCES AND USES THEREOF

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/757,669A

; FILING DATE:

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 45185-F

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3207 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-757-669A-20

Query Match 3.7%; Score 58; DB 3; Length 3207;

Best Local Similarity 52.5%; Pred. No. 0.0011;

Matches 127; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

Qy 1317 CGTCGCTTTCACGCTGCTCATGGCCATTATGCGCGCACTTCGCGCTGCTCATGGGCGCT 1376

Db 20703 CATGCCCGCGTCTCTCTGCT 20762

Qy 1377 CACGGCAGCTCACGGGCGCGGCGCTCTGTTTCTTGTCGCCAGGCTCTTTTACCTGGC 1436

Db 20763 CATGTCTCTCATCATGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20822

Qy 1437 CTTGCTCTGGCGCAAGTGTGTGGACCAAGTCTTCTTGAGGTGCGCATCTTGTCTCAT 1496

Db 20823 CTTCTCTGCT 20882

Qy 1497 CGGCGGCACTGTCAGGCGTGTCCGGCTTCGTGCACTCCCTTCGAGGGCTCATCGAAGCCTA 1556

Db 20883 CTTCTCTGCT 20942

Qy 1557 CC 1558

Db 20943 CC 20944

RESULT 15

US-09-230-371A-20

Qy	1497	CGGGGATCTGCAGCGTGTCCGGCTTCGTGCACTCCCTCGAGGGCTCATCGAAGCCTA	1556
Db	1114	ÇTCCTCGTCACTCCTCTCGTCAFCCTCGTCACTCCTCGTCACTCCTCTCGTCAT	1055
Qy	1557	CC	1558
Db	1054	CC	1053

```

RESULT 10
US-09-298-568-1/c
; Sequence 1, Application US/09298568
; Patent No. 6323792
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballestas, Mary E.
; APPLICANT: Kaye, Kenneth M.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/298,568
; CURRENT FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,422
; EARLIER FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-1

```

	Query Match	3.7%;	Score 58;	DB 3;	Length 3489;
	Best local Similarity	52.5%;	Pred. No. 0.00055;		
	Matches 127;	Conservative	0;	Mismatches 115;	Indels 0;
	Gaps	0;			
Qy	1317	CGTCGCTTTACGGCTGCTCATGGGCATTTATGTGCGGACATTTCGCGTGTCTCATGGGCGCT	1376		
Dp	1294	CATGCCCCCGTCTCTCTGTCCTCTTCTTGTCCTCTCTCTCATCTCTCTCGTCTCT	1235		
Qy	1377	CACGGGAGGCTCATAGGCGCGCGGCTCTGTTTCTGTGTGCCAGCTCTTTTTCACCTGTGG	1436		
Dp	1234	CATTGTCTCATCATCGTCATCTCTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1175		
Qy	1437	CTTGCTCTGGCGGAAGTGTGTGTGGACCAAGTCTTCTTGACCTCGGCATCTTCTGTCAT	1496		
Dp	1174	CTCTCTGTGTCCTCTCTCGTCATCTCTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT	1115		
Qy	1497	CGGCGGCATCTGCAGCGTGTCCGGCTCTGTGCATCTCCCTCGAGGGCGCTCATCGAAGCCTA	1556		
Dp	1114	CTCTCTGTGCATCTCTCTGTGCATCTCTCTGTGCATCTCTCTGTGCATCTCTCTGTGCAT	1055		

```

RESULT 11
US-09-410-399-1/c
; Sequence 1, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; APPLICANT: Cotter, Murray A.
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
; TITLE OF INVENTION: to Genomic Host DNA
; FILE REFERENCE: UM-03778
; CURRENT APPLICATION NUMBER: US/09/410.399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3489

```

```
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-410-399-1

Query Match          3.7%; Score 58; DB 4; Length 3489;
Best Local Similarity 52.5%; Pred. No. 0.00055;
Matches 127; Conservative 0; Mismatches 115; Indels 0; Gaps 0:

Qy      1317 CGTGGCTTTACGCTGCTCATGGCCATTTATGTGCCGCACTTTCGGCGTGCTCATGGGCCT 1376
         |||||
Db      1294 CATGCCCCCGCTCCTCGTCTCTCTTCTGTCTCTCTCTCGTCATCTCCTCGTCTCT 1235

Qy      1377 CACGGCAGCTCACGGGGCGCGGCTCTGTGTTCTTGTGCCCCAGGCTCTTTCACTGCG 1436
         |||||
Db      1234 CATTTGCTCTCATCATCGTATCGTCTCTCGTCTCTCTCTCTCTCTCTCTCGTCTCT 1175

Qy      1437 CTGTGCTCTGGCGCAAGCTGCTGTGGACCAAGTCTTTTCGACGTCGCCATCTTCGTCTAT 1496
         |||||
Db      1174 CCTCTCTGTGCTCTCTCTCGTCTATCTCTCTGTGTCATCTCTCTCGTCATCTCTCTGTCTAT 1115

Qy      1497 CGCGCGGCATCTGCAGCGGTCTCGGGTTCGTGCACTCCCTCGAGGGCGCTCATCGAAGCCTTA 1556
         |||||
Db      1114 CCTCTCTGTGTCATCTCTCTCGTCTATCTCTCTCGTCATCTCTCTCTCGTCTCTCTCGTCTAT 1055

Qy      1557 CC 1558
         ==
Db      1054 CC 1053
```

RESULT 12
US-09-894-273-1/c
; Sequence 1, Application US/09894273
; Patent No. 6756203
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballestar, Mary E.
; APPLICANT: Kaye, Kenneth M.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/894,273
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/109,422
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-1

Query Match	3.7%	Score 58	DB 4	Length 3489
Best Local Similarity	52.5%	Pred. NO. 0.00055		
Matches 127	Conservative	0	Mismatches 115	Indels 0
Gaps	0			
Qy	1317	CGTCGCTTTACACGCTGTCATGCGCCATTTATGFCGCGACATTGCGCGCTGCTCATGGCGCT	1376	
Db	1294	CATGCCCCGGTCCTCTCGTCTCTCTTCTTGTCCTCTCTCGTCATCTCTCTCGTCT	1235	
Qy	1377	CACGGCAGCCTCACGGGCGCGGCTCTGTTTCTTGTCGCCAGCGCTCTTTACACCTGG	1436	
Db	1234	CATTGTCCTCATCATGTCTATCTCTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCT	1175	
Qy	1437	CCTGCTCTGGCGAAGTGCTGTGGCACCAAGTCTTTTGGAGCTGCCCATCTTGCTCAT	1496	
Db	1174	CCTCTCGTCTCTCTCTCGTCATCTCTCTCGTCATCTCTCTCGTCATCTCTCTCGTCT	1115	
Qy	1497	CGGCGGCATCTGAGGGTGTCGGGCTTCGTGCACTCCCTCGAGGGCTCATCGAAGCCTA	1556	
Db	1114	CCTCCTCGTCATCTCTCTCGTCATCTCTCTCTCGTCATCTCTCTCTCTCTCTCTCT	1055	
Qy	1557	CC 1558		


```
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2057
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-2057

Query Match          7.0%; Score 110.8; DB 4; Length 981;
Best Local Similarity 53.8%; Pred. No. 2.7e-16;
Matches 229; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1138 GCGGTGTCACATCTTTCTGGTGGCCAGGCGCTTGTCTATCTCTCTGCAATCTTT 1197
DQ 2 GGCATGGTGAATCTTCTCTGGTGCATCAGAGCCCTTCTGAGTACCCCTTGCCCTACTAT 61

QY 1198 GCGGTGTCAGGTGTCGAGAAGTCTCTTTCCAGGAAGGCGCGCCCTTTTCCCG 1257
DQ 62 GTCGCTGTGAGTCTCGAACGCAACTTCTTTAGAGGTCCGCAAGCAAAATTTCCA 121

QY 1258 GCTGTCTACAGGGGCGAGCGGCGCTGAAGTCTCTGGGGCTGACGCTGGCTGGCGCTC 1317
DQ 122 ACCATCTGGAATCTGACGCGGGAATCTGAAGGTTTGGGGTCTGGGCTTCGGTGTGCGGCTC 181

QY 1318 GTGCTTTACGCTGCTCATGGCCATTTATGTGCGGCACTTCTGGGCTGCTCATGGGCTC 1377
DQ 182 ATTGATCCACCATCTATGATGGCCATTTTCACTCCCACTTCTCATCTCATGGGTTTC 241

QY 1378 ACCGCGACCTTCACGGGCGCGGCTCTGTCTTTCTGTGCGGCACTCTTTCACCTGGCG 1437
DQ 242 ATTGGAGCTTCACTGGCACAATGCTCAGCTTTATCTGGGCTTGTATTTCCACATCAAG 301

QY 1438 CTGCTCTGGGCAAGCTGCTGTGCAACCAAGTCTTCTTCGACGTGCGCATCTTCTGTCATC 1497
DQ 302 ATCAAGGGGCACTGCTCGATCAGAAGGAATAAGCAAGACTACCTCATCTTGGGCTC 361

QY 1498 GCGGCATCTGACGCTGTCGGGCTTGTGCACTCCCTTCGAGGGGCTCATCGAAGCCTTAC 1557
DQ 362 GCGTGTCTTTCGGGTTTATGTTATCTACGATTCGGGCAATGCCCTGATTAATGCAATTT 421

QY 1558 CGAACC 1563
DQ 422 GAAATC 427

RESULT 7
US-09-249-585A-4
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Query Match          4.4%; Score 68.8; DB 4; Length 1926;
Best Local Similarity 45.8%; Pred. No. 1.5e-06;
Matches 238; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

QY 1030 CACATCGCAGCGCTGCGTCTCAAGGGGCTCTTTCGGGCTCGTGCGCTACTCATCTGGGCG 1089
DQ 269 CTCGTCTCTGCTCTCGCCCTCCCGTCTCTGCTCTCTCCCGTCTCTGCTCTCTCTCCCGCTC 328

QY 1090 GACGAGACCAAGAGGTGATCAGGATAACCTGCGCGGCTCATCGGCGCGGTGCTCAAC 1149
DQ 329 CTCGTCTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCTCTCTCTCTCCCG 388

QY 1150 ATCTTTCTGGTGGCCAAAGCGCTGTGCTTCTATCTCTGCGCATCTTTTTCGGCTGTCGAG 1209
DQ 389 GTCTCTGCTCTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCTCTCTCTCTCTCTCTCTCT 448

QY 1210 GTGCTGGAGAGTCTGCTCTTTCAGGAAGCGACGCGCGCTTTTTCGGGCTCTGTACAGC 1269
DQ 449 GTCTCTCTCCCGTCTCTCCCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 508

QY 1270 GCGCAGCGGCGCTGAAAGTCTCTGGGGGTGACGCTGCGGCTGCGGCTGCTCTCTCTCTCTCT 1329
```

```
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2057
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-2057

Query Match          7.0%; Score 110.8; DB 4; Length 981;
Best Local Similarity 53.8%; Pred. No. 2.7e-16;
Matches 229; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1138 GCGGTGTCACATCTTTCTGGTGGCCAGGCGCTTGTCTATCTCTCTGCAATCTTT 1197
DQ 2 GGCATGGTGAATCTTCTCTGGTGCATCAGAGCCCTTCTGAGTACCCCTTGCCCTACTAT 61

QY 1198 GCGGTGTCAGGTGTCGAGAAGTCTCTTTCCAGGAAGGCGCGCCCTTTTCCCG 1257
DQ 62 GTCGCTGTGAGTCTCGAACGCAACTTCTTTAGAGGTCCGCAAGCAAAATTTCCA 121

QY 1258 GCTGTCTACAGGGGCGAGCGGCGCTGAAGTCTCTGGGGCTGACGCTGGCTGGCGCTC 1317
DQ 122 ACCATCTGGAATCTGACGCGGGAATCTGAAGGTTTGGGGTCTGGGCTTCGGTGTGCGGCTC 181

QY 1318 GTGCTTTACGCTGCTCATGGCCATTTATGTGCGGCACTTCTGGGCTGCTCATGGGCTC 1377
DQ 182 ATTGATCCACCATCTATGATGGCCATTTTCACTCCCACTTCTCATCTCATGGGTTTC 241

QY 1378 ACCGCGACCTTCACGGGCGCGGCTCTGTCTTTCTGTGCGGCACTCTTTCACCTGGCG 1437
DQ 242 ATTGGAGCTTCACTGGCACAATGCTCAGCTTTATCTGGGCTTGTATTTCCACATCAAG 301

QY 1438 CTGCTCTGGGCAAGCTGCTGTGCAACCAAGTCTTCTTCGACGTGCGCATCTTCTGTCATC 1497
DQ 302 ATCAAGGGGCACTGCTCGATCAGAAGGAATAAGCAAGACTACCTCATCTTGGGCTC 361

QY 1498 GCGGCATCTGACGCTGTCGGGCTTGTGCACTCCCTTCGAGGGGCTCATCGAAGCCTTAC 1557
DQ 362 GCGTGTCTTTCGGGTTTATGTTATCTACGATTCGGGCAATGCCCTGATTAATGCAATTT 421

QY 1558 CGAACC 1563
DQ 422 GAAATC 427

RESULT 6
US-09-270-767-17339
; Sequence 17339, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17339
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-17339

Query Match          7.0%; Score 110.8; DB 4; Length 981;
Best Local Similarity 53.8%; Pred. No. 2.7e-16;
Matches 229; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1138 GCGGTGTCACATCTTTCTGGTGGCCAGGCGCTTGTCTATCTCTCTGCAATCTTT 1197
DQ 2 GGCATGGTGAATCTTCTCTGGTGCATCAGAGCCCTTCTGAGTACCCCTTGCCCTACTAT 61
```

```
RESULT 3
US-09-270-767-15610/c
; Sequence 15610, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15610
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15610
Query Match      8.8%; Score 139.4; DB 4; Length 448;
Best Local Similarity 60.0%; Pred. No. 5.6e-23;
Matches 274; Conservative 0; Mismatches 171; Indels 12; Gaps 2;

Qy 369 GAACGTGACCAACCCATCCAGGCATGTTGCTGCTGGGCTACCCCTACGCCATCCCTGCA 428
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 448 GAACGTGACCAATGCCATTCAGGCATGTTTCATCATCTTCGCCGCCGTTGTGCTGTACACGG 488
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 429 CGGCGGCTACCTGGGTTGTTCTCATCATCTTCGCCGCCGTTGTGCTGTACACGG 488
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 388 TGGCGGCTATTGGGCCATCGGGCCATGGTGGGAATAGCACATCTGCTGTATACGGG 329
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 489 CAAGATCTCATCGCGTCCCTGTACGA---GGAGAAATGAAGACGGCGAGTGGTGGCGGT 545
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 328 CAAGTCTGCTGTCAGTGTCTATACGAACCGGATCCGCGACGGGTCAATGTTGGCGAGT 269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 546 GCGGACTCGTACGTGGCCATAGCAACGCTGTGCGGCCCGCGCTTCCCAACGCTGGG 605
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 268 GCGCGACAGTTATGCGCCATAGCAAGGTATGTTTGT-----GTCCAAAGTTGGG 218
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 606 GCGCGAGTGGTGAACCTAGCGGAGATCATCGAGCTGGTATGACGTGATCCTGTACGT 665
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 217 CGCCCGGCGAGTCAGCATTCGCCAGTTCATCGAGTTCATGACCTGTCATCTGTACGT 158
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 666 GGTGTGAGTGGCAACCTCATGTACAACAGCTTCCCGGGCTGCCGCTGTCGAGAGTC 725
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 157 GGTGTGCGCGATCTATTGGCGGAAGTATCCACAGGCTGTTTCACTCGCGATC 98
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 726 CTGCTCCATTATCGCCACGGCCGCTGTGCTGCTGCTGCTGCTTCTTAAGAACCTCAAGGC 785
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 97 CTGGATGCTATTCTGGGGCATATTCTCTGCTGCCATGGGATTCCTTAAGTCACTGAAAT 38
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 786 CGTGTCCAAGTTCAGTCTGCTGTGTCATCTGGGCCAC 822
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 37 GGTATCAACGCTCTCGTTCTGGTGCACAATGTGCGAC 1
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-09-178-093B-3
; Sequence 3, Application US/09178093B
; Patent No. 6660846
; GENERAL INFORMATION:
; APPLICANT: Robert H. Edwards
; APPLICANT: Richard J. Reimer
; APPLICANT: Steve L. McIntire
; APPLICANT: Erik M. Jorgenson
; APPLICANT: Kim Schuske
; TITLE OF INVENTION: Vesicular Amino Acid Transported
; FILE REFERENCE: Composition and Method
; CURRENT APPLICATION NUMBER: US/09/178,093B
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/063,012
; PRIOR FILING DATE: 1997-10-23
```

```
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; NAME/KEY: misc feature
; LOCATION: (1)-(1586)
; OTHER INFORMATION: UNC-47 cDNA
US-09-178-093B-3
Query Match      8.7%; Score 137.2; DB 4; Length 1586;
Best Local Similarity 50.6%; Pred. No. 2.7e-22;
Matches 359; Conservative 0; Mismatches 348; Indels 3; Gaps 1;

Qy 727 TGGTCCATTATCGCCACGGCGGTGCTGCTGCTGGCCCTTCTTAAAGAACCTCAAGGCC 786
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 610 TGGATGATGATTACCTCAGCATCTTTACTAACGTGCTCAATTTCTTGATGATCTACAAT 669
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 787 GTCTCCAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 670 GTGCTCGTGTGTCATTTTCAATGCAATATCTCATTTGATTTGTCATCTGATCATGCTC 729
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 847 GCCTACTGCTATCGCGGGCGCGACTGGGCTGGGAGAGTCAAGTTCTACATCGAC 906
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 730 CTTTACTGCTGTCATTCGCTCACAATGCTCTTCTCAACGATCACATTTTCATTGAAT 789
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 907 GTCAAGAAGTTCCTCATCTCCATTGGCATCATCGTGTTCAGCTACAGCTCTCAGATCTTC 966
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 790 ATCAACACTCTTCGACAAATTTGGAAATGGTTGTTTTCGGCTACACATCTCATATATTC 849
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 967 CTCGCTTCGCTGAGGCAATATGAGAGCCAGGAGTTCCACTGCATGATGAACCTGG 1026
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 850 CTTTCAAAATTTAGAAGGAAATATGAAAAATCTCTGCTCAATTCACAGTAATGTTAAATGG 909
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1027 ACCCATCATCGACCTCGCTGCTCAAGGGCTCTTCGCGCTCGCTGCTGCTGCTGCTGCTG 1086
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 910 TCACACATCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1087 GCGCAGCAGACCAAGAGGTTCATCAGGATAAATCTGCC---CGGCTCCATCCGCGCGGTG 1143
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 970 GGAGAGCTTACACAGAGGAAATTTCAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1029
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1144 GTCAACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1203
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1030 GTGAACCTGATTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1089
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1204 GTCGAGGTGCTGAGAGTCTGCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1263
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1090 GTTCAACTTTTGAAGAACAAATTTGTTCTTGGATATCTTCAAGACCAATTTCAAGTTGT 1149
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1264 TACAGCGCGCAGCGGCGCTGAAAGTCTTGGGGGCTGACGCTGCGCTGCGGCTGCTGCTG 1323
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1150 TATTCACCGGATAAATCTTTACGTGAATGGCGGTTACTTTAAGAAATTTCTAGTGTCT 1209
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1324 TTCACGCTGCTATGCGCAATTTATGCGCGCACTTCGCGCTGCTGCTGCTGCTGCTGCTG 1383
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1210 TTCACACTTTTCGTTGCTATTCAGTTCCTATATTTGGTAGAGTTGATGGGATAGTTGGA 1269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1384 AGCTCAGCGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1433
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1270 AATATTACAGGAACAATGTTTATCATTTATCTGCGCGGCACTATTCCACCT 1319
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-09-270-767-2057
; Sequence 2057, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 18:34:45 ; Search time 159 Seconds
(without alignments)
7054.244 Million cell updates/sec

Title: US-09-940-919-1
Perfect score: 1578
Sequence: 1 atggccacctgtctcgag.....gaaccaacggaggactag 1578

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1352.2	85.7	2392	4	US-09-178-093B-4
2	139.4	8.8	448	4	US-09-270-767-328
3	139.4	8.8	448	4	US-09-270-767-15610
4	137.2	8.7	1586	4	US-09-178-093B-3
5	110.8	7.0	981	4	US-09-270-767-2057
6	110.8	7.0	981	4	US-09-270-767-17339
7	68.8	4.4	1926	4	US-09-249-585A-4
8	68.8	4.4	1931	2	US-09-130-114-2
9	58	3.7	3489	2	US-08-728-323A-1
10	58	3.7	3489	3	US-09-298-568-1
11	58	3.7	3489	4	US-09-410-399-1
12	58	3.7	3489	4	US-09-894-273-1
13	58	3.7	32207	2	US-08-770-379-20
14	58	3.7	32207	3	US-08-757-669A-20
15	58	3.7	32207	3	US-09-230-371A-20
16	55.4	3.5	885	4	US-09-252-991A-2480
17	55.4	3.5	921	4	US-09-252-991A-2561
18	55.4	3.5	2532	4	US-09-252-991A-2356
19	53.8	3.4	948	4	US-09-252-991A-14810
20	53.8	3.4	1428	4	US-09-252-991A-14802
21	53.8	3.4	1668	4	US-09-252-991A-14950
22	53.8	3.4	1926	4	US-09-249-585A-2
23	53.8	3.4	1926	4	US-09-410-399-3
24	53.8	3.4	2580	3	US-09-050-863-2
25	53.8	3.4	2580	3	US-09-359-081-2
26	53.8	3.4	5452	2	US-09-130-114-1
27	53.8	3.4	8705	4	US-09-647-344A-14

C 28	53.8	3.4	9600	3	US-08-910-647-1	Sequence 1, Appli
C 29	53.8	3.4	9600	4	US-09-620-925-1	Sequence 1, Appli
C 30	53.8	3.4	10596	1	US-07-884-811-15	Sequence 15, Appli
C 31	53.8	3.4	10596	1	US-07-885-971-15	Sequence 15, Appli
C 32	53.8	3.4	10596	1	US-08-087-783A-15	Sequence 15, Appli
C 33	53.8	3.4	10596	1	US-08-194-088B-15	Sequence 15, Appli
C 34	53.8	3.4	10596	2	US-08-194-087-15	Sequence 15, Appli
C 35	53.8	3.4	10596	5	PCT-US93-04648-15	Sequence 15, Appli
C 36	53.8	3.4	16080	4	US-09-724-566A-48	Sequence 48, Appli
C 37	52.8	3.3	2571	4	US-09-252-991A-9022	Sequence 9022, Ap
C 38	52.8	3.3	3150	4	US-09-252-991A-8765	Sequence 8765, Ap
C 39	52.8	3.3	3450	4	US-09-252-991A-8986	Sequence 8986, Ap
C 40	52.6	3.3	7218	1	US-08-232-463-14	Sequence 14, Appli
C 41	50.8	3.2	1578	4	US-09-252-991A-10972	Sequence 10972, A
C 42	50.8	3.2	2547	4	US-09-252-991A-11271	Sequence 11271, A
C 43	50.8	3.2	2964	4	US-09-252-991A-10912	Sequence 10912, A
C 44	50.2	3.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 45	50.2	3.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-178-093B-4
; Sequence 4, Application US/09178093B
; Patent No. 6660846
; GENERAL INFORMATION:
; APPLICANT: Robert H. Edwards
; APPLICANT: Richard J. Reimer
; APPLICANT: Steve L. McIntire
; APPLICANT: Erik M. Jorgenson
; APPLICANT: Kim Schuske
; TITLE OF INVENTION: Vesicular Amino Acid Transported
; FILE REFERENCE: 2002-0005.30
; CURRENT APPLICATION NUMBER: US/09/178,093B
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/063,012
; PRIOR FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 2392
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2392)
; OTHER INFORMATION: RUNC-47 cDNA
US-09-178-093B-4

Query Match 85.7%; Score 1352.2; DB 4; Length 2392;

Best Local Similarity 90.8%; Pred. No. 2.1e-301;

Matches 1433; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY	1	ATGCCACCTTGTCTCGCAGCAGCTGTCCACGTCGCCACGTCCTGTGTCCACAGTCC	60
DB	87	ATGCCACCTTGTCTCGCAGCAGCTGTCCACGTCGCCACCTCTGTGTCCACAGTCC	146
QY	61	CAGGCCAAGATGAGCGGCATGTTCCGCGAGTGGGTTTTTCAGCGGCCACGGATGAGGAG	120
DB	147	CAGGCCAAGTGTGAGCGGCATGTTCCGCGAGTGGGTTTTTCAGCGGCCACGGATGAGGAG	206
QY	121	GGCGTGGGCTTCGCGCATTTGCCAGCACCTTCGATTTAGCACCGCCAGGCGCTTCGAGATG	180
DB	207	GGCGTGGGCTTCGCGCATTCGCGCATCTTCGATTTAGCACCGCCAGGCGCTTCGAGATG	266
QY	181	GACATCTCTGAAGCCGAGGAGAGCCCTCGGGGACGAGGGCGCTGAAGGCCCGCTCGAG	240
DB	267	GACATCTCTGAAGCCGAGGAGAGCCCTCGGGGACGAGGGCGCTTCGCGCGCTCGAG	326
QY	241	GGAGACATCCATTATTCAGCGGAGCGAGCTCTCTTCGCGGCCCTCCGCTCCAAGGAC	300

Qy

432

CGGCTACTCGGGTGTCTTCATCATCTTCGCGCGTGTGTGTGCTGTACACCGCAA

491

Db

1579

CGGCTATTGGCCATCGTGGCCATGTGGNATAGCACACATCTCTGTCTATACGGCAA

1638

Qy

492

GATCCTCATCGCGTGCCTGTACGA---GGAGAAATGAAGACGGCGAGTGTGTGCGGTGCG

548

Db

1639

GGTCTGTGTCAGTGTCTATACGAACCGGATCCGGCAACGGACAATGTGCGAGTGG

1698

Qy

549

GGACTCGTACGTGGCCATAGCCACCGCTGTGCGCCCGCGCTTCCCAACGCTGGGGCG

608

Db

1699

CGATAGTTATGTGGCCATAGCCAAAGTATGTTTTG-----GTCCAAAATTGGGCG

1749

Qy

609

CCGAGTGTGAACGTAGCGAGATCATCGAGCTGTGATGACGTGCATCCTGTACGTGGT

668

Db

1750

CCGGGAGTCAGCATGTCGCCAGCTTTCGAGCTTCTGATGACCTGCATCCTGTACGTGGT

1809

Qy

669

GGTGAATGGCAACCTCATGTACAACAGCTTCCGGGGCTGCCCGTGTGCGAGAAGTCCCTG

728

Db

1810

GGTGTGCGGATCTATTGGCCGGAACGTATCCACAGGGCTGTTGCACTCGCGATCCTG

1869

Qy

729

GTCCTATTGCGCACGGCGGTGCTGTGCTTGGCCCTTCTTAAAGAACCTCAAGGCCGT

788

Db

1870

GATGCTATTGCTGGGCATATTCTCTGCTGCCCATGGGATCTTAAAGTCACTGAAAATGGT

1929

Qy

789

GTCCAAGTTCAGTCTGTGTGCACTCTGCGCCACTTCGTGCATCAATATCCTGTGTCATGC

848

Db

1930

ATCAACGCTCTGTTCTGTGTGCACAATGTGCAATCGTANTAAACCCGTGATCTGGG

1989

Qy

849

CTACTGTCTATCGCGGGCGCGCACTGGGCGCTGGGAGAAGTCAAGTTCATACATCGACGT

908

Db

1990

CTATTGCTCTGTCAGATTGGCGATTGGGGTTGGTCCAAAGTTCGATTAGCATCGACAT

2049

Qy

909

CAAGAACTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACACGTCTCAGATCTTCCT

968

Db

2050

GGAGAACTTTCGATCTCACTGGGCGTCATCGTCTTCTCGTACACCTCCAGATCTTTCT

2109

Qy

969

GCCTTCGCTGGAGGCATATGCAGCAGCCAGCGAGTCCACTGCATGATGACTGGAC

1028

Db

2110

GCCACGCTCGAGGGCAACATGATCGATCGTCCAAATCAACTGGATGCTGGACTGGTC

2169

Qy

1029

GCATATCGACCTCGCTGTCTCAAGGGCTCTTCGCGCTGTGCTACCTCACCTGGGC

1088

Db

2170

TCACATAGCGCTGCTGTGTTTAAAGCCGGATTGGGTATATCTGCTTTCTGACCTTCCA

2229

Qy

1089

CGAGGACCAAGAGGTCAATCAGGATACCTGCG---CCGGTCCATCCGCGCCGTGGT

1145

Db

2230

GAACGATACACAGCAGGTGATTACCAACAATCTGCACTCGCAGGGCTTTAAGGGCATGGT

2289

Qy

1146

CAACATCTTCTGTGGCCAGGCGTGTGCTCTATCTCTGCCATTCTTTGCCGCTGT

1205

Db

2290

GAACTTCTTCTGTCTATCAAGGCCCTTCTGAGTACCCCTTGCCCTACTATGCTGCCCTG

2349

Qy

1206

CGAGGTCTGGAGAACTCGCTCTTCCAGGAAGCAGCGCGCTTTTCCCGGCGCTGCTA

1265

Db

2350

TGAGTCTCTGACGCAACTTCTTTAGAGTCCGCCAAAGACCAATTTCCACCATCTG

2409

Qy

1266

CAGCGGCGACGGGCGCTGAAGTCTGCGGGCTGAAGCTGCGTGGCGCTCGTGTGCTT

1325

Db

2410

GAATCTGGACGGGAACTGAAGGTTTGGGGTCTGGGCTTCCGTGTCCGGCGCTATTGTATC

2469

Qy

1326

CAGCTGCTCATGGCCATTATGTGCGCACTTCGCGCTGCTCATGGGCGCTCACCGGCG

1385

Db

2470

CACCATACTGATGGCCATTTTCATTTCCCACTTCTCCATTTCTGATGGGTTTCATTTGGCAG

2529

Qy

1386

CCTCACGGGCGCGGCTCTGTTTCTTGTGCGCCAGGCTCTTTTCCCTGGCGCTGCTCTG

1445

Db

2530

CTTCACTGGCAATGCTCAGCTTATCTGCGCTTGTCTATTTCCACATCAAGATCAAGG

2589

Qy

1446

GCGCAAGCTGCTGTGGCAACAAAGTCTTCTTCAGAGTGGCCATCTTGTCTATCGGCGG

1505

Db

2590

GCATCTGCTCGATCAGAAGGAATAGCCAAAGACTACCTCATCATTTGGGCTCGGCGTCT

2649

Search completed: November 3, 2004, 19:15:35
Job time : 760 secs


```
XX AAS81399;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #17203.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG17212.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 17203; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (II) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activities. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1095 BP; 240 A; 341 C; 338 G; 176 T; 0 U; 0 Other;
XX
XX Query Match 24.5%; Score 387.2; DB 5; Length 1095;
XX Best Local Similarity 99.2%; Pred. No. 2.9e-69;
XX Matches 389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 ATGCCACCTTGCTCCGAGCAAGCTGTCCAAAGTGGCCACGTCGTCGTCACAAAGTCC 60
XX
XX 220 ATGGCCACCTTGCTCCGAGCAAGCTGTCCAAAGTGGCCACGTCGTCGTCACAAAGTCC 279
XX
XX 61 CAGGCCAAGATGAGCGCATGTTCCGAGCATGGTTTTCAGGCGGCGCAGGATGAGG 120
XX
XX 280 CAGGCCAAGATGAGCGCATGTTCCGAGCATGGTTTTCAGGCGGCGCAGGATGAGG 339
XX
XX 121 GCGGTGGGCTTCGCGCATTTGCGAGCATCTCGACTTTTGAGCACCGCCAGGCGCTGCAGATG 180
XX
XX 340 GCGGTGGGCTTCGCGCATTTGCGAGCATCTCGACTTTTGAGCACCGCCAGGCGCTGCAGATG 399
XX
XX 181 GACATCTGAAAGCCGAGGAGAGCCCTTCGCGGAGACGAGGCGCTGAAGCGCCGTCGAG 240
XX
XX 400 GACATCTGAAAGCCGAGGAGAGCCCTTCGCGGAGACGAGGCGCTGAAGCGCACGTCGAG 459
XX
XX 241 GGAGACATCATTTATCAGCGAGGAGGAGGAGCTCTCTGCGGCCCTCCGGCTCCAAGGAC 300
XX
XX 460 GGAGACATCATTTATCAGCGAGGAGGAGGAGCTCTCTGCGGCCCTCCGGCTCCAAGGAC 519
XX
XX 301 CAGGTGGGAGTGGTGGCGAATTTCGGGGGCCACGACAAAGCCCAAAATCAGGCGCTGGGAG 360
XX
XX 520 CAGGTGGGAGTGGTGGCGAATTTCGGGGGCCACGACAAAGCCCAAAATCAGGCGCTGGGAG 579
XX
XX 361 GCAGGCTGGAAACGTGACCAACGCCATCCAGGG 392
XX
XX 580 GCAGGCTGGAAACGTGACCAACGCCATCCAGTG 611
XX
XX RESULT 14
XX ABL14045
XX ID ABL14045 standard; cDNA; 1650 BP.
XX AC ABL14045;
XX XX
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 36617.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX XX
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX XX
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX XX
XX DR WPI; 2001-656860/75.
XX DR P-PSDB; ABB69942.
XX XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX XX
XX PS Claim 1; SEQ ID NO 36617; 21pp + Sequence Listing; English.
XX XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1650 BP; 375 A; 438 C; 445 G; 392 T; 0 U; 0 Other;
XX
XX Query Match 23.2%; Score 366; DB 4; Length 1650;
XX Best Local Similarity 57.6%; Pred. No. 6.4e-65;
XX Matches 720; Conservative 0; Mismatches 515; Indels 15; Gaps 3;
```

XX Mus sp.
OS Synthetic.
XX WO200136632-A2.
XX 25-MAY-2001.
XX 17-NOV-2000; 2000WO-IL000766.
XX 17-NOV-1999; 99IL-00132978.
XX 10-DEC-1999; 99IL-00133455.
XX (COMP-) COMPUGEN LTD.
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX WPI; 2001-336004/35.
XX P-PSDB; AAU02987.
XX Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX Claim 1; Page 374; 519pp; English.
XX The sequence represents a DNA encoding an angiotensin converting enzyme
CC splice variant (ACEV) polypeptide. The polypeptides of the invention
CC include variants of granulocyte colony stimulating factor receptor,
CC glucagon, interleukin 6, platelet-derived endothelial cell growth factor,
CC cyclin-dependent kinase inhibitor 1C, cellular tumour antigen p53, and
CC vasoactive intestinal polypeptide receptor 2. The polypeptides and their
CC associated nucleic acids are useful for identification of variant
CC sequences and detection of candidate compounds capable of binding the
CC molecules. The sequences of the invention can be used in the treatment
CC as and diagnosis of various disorders including cardiovascular diseases such
CC as arteriosclerosis, myocardial infarction and coronary arterial
CC thrombosis, renal diseases such as diabetic nephropathy, muscular
CC diseases such as hypertrophy, immune disorders such as immune complex
CC nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic
CC pulmonary granulomatous diseases such as asbestosis and vascular
CC pathologies involving an endothelial abnormality such as deep vein
CC thrombosis
XX Sequence 1723 BP; 364 A; 506 C; 534 G; 319 T; 0 U; 0 Other;
Query Match 72.2%; Score 1140; DB 5; Length 1723;
Best Local Similarity 91.4%; Pred. No. 2.1e-222;
Matches 1232; Conservative 0; Mismatches 110; Indels 6; Gaps 2;
QY 1 ATGGCCACCTTGTCTCCGACGAGCTGTCCACGTGGCCACGTCGCTGTCCACAGTCC 60
DB 1517 ATGGCCACCTTGTCTCCGACGAGCTGTCCACGTGGCCACCTCGTGTCCACAGTCC 1458
QY 61 CAGGCCAAGATGAGCGGCATGTTTCGCCAGGATGGGTTTTCAGCGGCCACCGATGAGGAG 120
DB 1457 CAGGCCAAGTGAAGCGGCATGTTTCAGGATGGGTTTTCAGCGGCCACCGATGAGGAA 1398
QY 121 GCGGTGGGCTTCGCGATTTGCGACGACCTCGACTTTGAGACGCGCCAGGCGCTGCGAGATG 180
DB 1397 GCGGTGGGCTTCGCGACCTGCGACGATCTCGACTTTGAGACGCGCCAGGCGCTGCGAGATG 1338
QY 181 GACATCTCAAGCCAGGAGGAGCCCTCGGGGACGAGGCGCTGAAAGCGCCGCTCGAG 240
DB 1337 GACATCTCAAGATCGGAAGCGGAGCCCTCGGAGACGAGGCGCGAGAACTCCCGTCGAG 1278
QY 241 GGAGACATCATTTATCAGCGAGGAGGAGCTCTCTCGCGCCCTCCCGGCTCCAGGAC 300
DB 1277 GGAGACATTTATTCAGC---GCGGCGGCGCTCTCTCGCACCTCTGCTGCTCCAGGAC 1221
QY 301 CAG---GTGGGAGTGTGGCGAATTTGGGGGGCCACGACAAAGCCCAAAATCAAGCGGTGG 357
DB 1220 CAGGCGGTGGAGTGTGGGAGTTGGGGGTTCAGCAAAACCCCAAGATCAAGCGGTGG 1161

QY 358 GAGGAGGCTGGAACGTCGACCAAGCCATCCAGGCGATGTTGCTGCTGGGCTTACCCTAC 417
DB 1160 GAAGCGGCTGGAACGTCGACCAAGCCATCCAGGCGATGTTGCTGCTGGGCTTACCCTAC 1101
QY 418 GCCATCTGCAAGCGGCTACCTCGGGTGTGTTCTCATCATCTTCGCGCGGCTGTGTGC 477
DB 1100 GCCATCTGCAAGCGGCTACCTCGGGTGTGTTCTCATCATCTTCGCGCGGCTGTGTGC 1041
QY 478 TGCTACACCGGCAAGATCTCATCGCGTGCCTGACGAGGAGAAATGAGAGCGGCGAGTG 537
DB 1040 TGCTACACCGGCAAGATCTCATCGCGTGCCTGACGAGGAGAAATGAGAGCGGCGAGTG 981
QY 538 GTGCGGTGCGGAGCTCGTACGTGCGCATAGCAACGCTGCTGCGCGCCCGCGCTTCCCA 597
DB 980 GTGCGGTGCGGAGCTCGTATGTGCGCATAGCAACGCTGCTGCGCGCTTCCCA 921
QY 598 ACCTGCGGCGCGGAGTGTGAAGTACGCGAGATCATCGAGCTGGTGAATGAGCGTGCATC 657
DB 920 ACCTGCGGCGCGGAGTGTGAAGTACGCGAGATCATCGAGCTGGTGAATGAGCGTGCATC 861
QY 658 CTGTACGTGTGTGAGTGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCCGTGTGC 717
DB 860 TTGTACGTGTGTGAGCGGCAACCTCATGTACACAGTTCCTCCCGGGGCTGCCCGTGTGC 801
QY 718 CAGAGTCTCTGCTCATTTATCGCCACGCGCTGCTGCTGCTGCGCTTCTTAAAGAAC 777
DB 800 CAGAGTCTCTGCTCATTTATCGCCACGCGCTGCTGCTGCTGCGCTTCTTAAAGAAC 741
QY 778 CTCAGGCGGTGTCAAGTTCAGTCTGTGTGCACTCTGTGGCCACTTCGTGTATCAATATC 837
DB 740 CTCAGGCGGTGTCAAGTTCAGTCTGTGTGCACTCTGTGGCCACTTCGTGTATCAATATC 681
QY 838 CTGTGTACAGCTACTGTCTATCGCGGCGCGGAGTCTGGGCTGGAGAGTCAAGTTC 897
DB 680 CTGTGTACAGCTACTGTCTATCGCGGCGCGGAGTCTGGGCTGGAGAGTCAAGTTC 621
QY 898 TACATCGAGCTCAAGAGTTCCTCATCTCCATTTGCGCATCATCGTGTTCAGCTTACAGTCT 957
DB 620 TACATCGAGCTCAAGAGTTCCTCATCTCCATTTGCGCATCATCGTGTTCAGCTTACAGTCT 561
QY 958 CAGATCTTTCTGCTGCTGAGGCGCAATATGACAGCGCCAGCGAGTTCACATGCAATG 1017
DB 560 CAGATCTTTCTGCTGCTGAGGCGCAATATGACAGCGCCAGCGAGTTCACATGCAATG 501
QY 1018 ATGAATCGACGACATCGACCTGCTGCTCAAGGCGCTTTCGCGCTGCTGCGCTTAC 1077
DB 500 ATGAATCGACGACATCGACCTGCTGCTCAAGGCGCTTTCGCGCTGCTGCGCTTAC 441
QY 1078 CTCACCTGGGCGGACGAGCAAGAGGCTCATACGCGATTAACCTGCGCGGCTTCCATCCGC 1137
DB 440 CTCACCTGGGCGGACGAGCAAGAGGCTCATACGCGATTAACCTGCGCGGCTTCCATCCGC 381
QY 1138 GCGGTGTCAACATCTTTCTGTGTGCCAAGGCGCTGTGTCTATCTCTGCGCATTTCTTT 1197
DB 380 GCGGTGTCAACATCTTTCTGTGTGCCAAGGCGCTGTGTCTATCTCTGCGCTTCTTT 321
QY 1198 GCGGTGTGAGGTGCTGAGAGTCTGCTTTCAGAGGCGAGCGCGGCTTTTCCCG 1257
DB 320 GCGGTGTGAGGTGCTGAGAGTCTGCTTTCAGAGGCGAGCGCGGCTTTTCCCG 261
QY 1258 GCGGTGTGAGGTGCTGAGAGTCTGCTTTCAGAGGCGAGCGCGGCTTTTCCCG 1317
DB 260 GCGGTGTGAGGTGCTGAGAGTCTGCTTTCAGAGGCGAGCGCGGCTTTTCCCG 201
QY 1318 GTGCTTTTACGCTGCTCATGCGCATTT 1345
DB 200 GTGCTTTTACGCTGCTCATGCGCATCT 173

RESULT 13
AAS81399
ID AAS81399 standard; cdNA; 1095 BP.

CC sequences and detection of candidate compounds capable of binding the
CC molecules. The sequences of the invention can be used in the treatment
CC and diagnosis of various disorders including cardiovascular diseases such
CC as arteriosclerosis, myocardial infarction and coronary arterial
CC thrombosis, renal diseases such as diabetic nephropathy, muscular
CC diseases such as hypertrophy, immune disorders such as immune complex
CC nephritis, multiple sclerosis, cancer, sarcoidosis, nonarctoidotic
CC pulmonary granulomatous diseases such as asbestosis and vascular
CC pathologies involving an endothelial abnormality such as deep vein
CC thrombosis
XX
SQ

Sequence 1648 BP; 278 A; 540 C; 480 G; 350 T; 0 U; 0 Other;

Query Match 85.2%; Score 1344.2; DB 5; Length 1648;
Best Local Similarity 90.5%; Pred. No. 5.8e-264;
Matches 1428; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Qy 1 ATGGCCACCTTGTCTCCGAGCAAGCTGTCCAACTGGCCAGCTCCGTGTCTCAACAAGTCC 60
Db |||||
Qy 42 ATGGCCACCTTGTCTCCGAGCAAGCTGTCCAACTGGCCAGCTCCGTGTCTCAACAAGTCC 101
Db |||||
Qy 61 CAGGCCAAGATGAGGGGATGTTCCGAGGATGGTTTTCAGGCGGCCACGGATGAGGAG 120
Db |||||
Qy 102 CAGGCCAAGATGAGGGGATGTTTCCAGGATGGGTTTTCAGGCGGCCACGGATGAGGAA 161
Db |||||
Qy 121 CGGTGGGCTTCCGCGATTGCGAGCACTCCACTTTGAGCACCCAGGGGCTTCAGATG 180
Db |||||
Qy 162 CGGTGGGCTTCCGCGATTGCGAGCACTCCACTTTGAGCACCCAGGGGCTTCAGATG 221
Db |||||
Qy 181 GACATCTCTGAAGCCGAGGAGAGCCCTGCGGGGACAGAGGCGCTGAAGCCGCCGTCGAG 240
Db |||||
Qy 222 GACATCTCTGAAGCCGAGGAGAGCCCTGCGGAGCAGGAGGCGCAGAGCTCCGTCGAG 281
Db |||||
Qy 241 GGAGACATTCATATTCAGGAGGAGGAGGAGCTCTTTCGCGCCCTCCGGCTCCAAAGAC 300
Db |||||
Qy 282 GGAGACATTCATATTCAGGAGGAGGAGGAGCTCTTTCGCGCCCTCCGGCTCCAAAGAC 341
Db |||||
Qy 301 CAGTGGGAGTGTGGCGAATTCGGGGGACACAGAGCCCAAAATCAAGCGCTGGAG 360
Db |||||
Qy 342 GCCGTGGAGTGTGGGGAGTTCGGGGGTTCACACAAACCAAGATCAAGGCGTGGGAA 401
Db |||||
Qy 361 CGAGGCTGGAGTGTGGAGCCATCCAGGCGATGTTGCTGCTGGGCTACCCCTACGCC 420
Db |||||
Qy 402 CGGGCTGGAGTGTGACAAATGCAATTCAGGCGATGTTGCTGCTGGGCTACCCCTACGCC 461
Db |||||
Qy 421 ATCTGTCAGCGCGCTACCTGGGTTGTTCTCATATCTTTCGCGCCGCTTGTGTGCTGC 480
Db |||||
Qy 462 ATCTTCCACGCGCGCTACCTGGGTTGTTCTCATCATCTTTCGCGCGAGTGTGTGCTGC 521
Db |||||
Qy 481 TACACCGGAGATCTCATCGCTGCTGTACGAGGAGATGAAGACGGGAGGTGGTG 540
Db |||||
Qy 522 TACACCGGAGATCTCATCGCTGCTGTACGAGGAGATGAAGACGGGAGGTGGTG 581
Db |||||
Qy 541 CGCGTGGGAGTCTGTAGTGGCCATAGCCAAACCCCTGCTGCGCCCGCGCTCCCAACG 600
Db |||||
Qy 582 CGCGTGGGAGTCTGTATGTGGCCATAGCTAAACCATGCTGCGCTCTCATATCCCAAC 641
Db |||||
Qy 601 CTGGGCGGCGAGTGTGTAAGATGAGCGAGATCATGAGCTGTGTGATGAGCTGCATCTTG 660
Db |||||
Qy 642 CTGGGCGGCGAGTGTGTAAGATGAGCGAGATCATGAGCTGTGTGATGAGCTGTATCTTG 701
Db |||||
Qy 661 TACGTGTGTGAGTGTGAGCACTTATGTAAGAGCTTCCCGGGCTGCGGTGTGCGAG 720
Db |||||
Qy 702 TACGTGTGTGAGTGTGAGCACTTATGTAAGAGCTTCCCGGGCTGCGGTGTGCGAG 761
Db |||||
Qy 721 AAGTCTGTGTGTAATTCAGCGCGCTGCTGCTGCTGCTGCTGCTTCCCTTAAGAACCTC 780
Db |||||
Qy 762 AAGTCTGTGTGTAATTCAGCGCGCTGCTGCTGCTGCTGCTGCTTCCCTTAAGAACCTC 821
Db |||||
Qy 781 AAGCGCGTGTCAAAGTTCAGTCTGCTGTGCACTCTGCGCCCACTTTCGTCATCAATATCTTG 840
Db |||||
Qy 822 AAGCGCGTGTCAAAGTTCAGTCTGCTGTGACGCTGCGCCCACTTTCGTCATCAATATCTTG 881
Db |||||

Qy 841 GTCATAGCTACTGTCTATTCGCGGCGCGAGCTGGGCTGGGAGAGGTCAAGTTCATC 900
Db |||||
Qy 882 GTCATAGCTACTGTCTATTCGCGGCGCGAGTATGGGCTGGGAGAGGTCAAGTTCATC 941
Db |||||
Qy 901 ATCGACGCTCAAGAGTTCCTCCATCTCCATTTGGCATCATCTGTTGTTTTCAGCTACACGTTCTCAG 960
Db |||||
Qy 942 ATCGACGCTCAAGAGTTCCTCCATCTCCATTTGGCATCATCTGTTGTTTTCAGCTACACGTTCTCAG 1001
Db |||||
Qy 961 ATCTTCTGCTTCTGCTGGAGGCAATATGACAGAGCCAGCGAGTTCCTACATGATGATG 1020
Db |||||
Qy 1002 ATCTTCTGCTTCTGCTGGAGGCAATATGACAGAGCCAGCGAGTTCCTACATGATGATG 1061
Db |||||
Qy 1021 AACTGGACGACATCGCAGCTGCTCAAGGGCTCTTTCGCGCTGCTGCGCTACCTC 1080
Db |||||
Qy 1062 AACTGGACGACATCGCAGCTGCTCAAGGGCTCTTTCGCGCTGCTGCGCTACCTC 1121
Db |||||
Qy 1081 ACCTGGGCGCCAGACGAGCAAGAGGCTCATCACGGATTAACCTTCCCGGCTTCCATCCGGGCC 1140
Db |||||
Qy 1122 ACCTGGGCGCCAGACGAGCAAGAGGCTCATCACGGATTAACCTTCCCGGCTTCCATCCGGGCC 1181
Db |||||
Qy 1141 GTGCTCAACATCTTCTGCTGGGCAAGGCGTGTGCTTATCTCTGCTTTCGCTTTCGCT 1200
Db |||||
Qy 1182 GTGCTCAACATCTTCTGCTGGGCAAGGCGTGTGCTTATCTGCTTTCGCTTTCGCT 1241
Db |||||
Qy 1201 GCTGTCAGGCTGTGGAGAGTCTCTTCCAGGAGGCGAGCGGCTTTCGCTTTCGCT 1260
Db |||||
Qy 1242 GCTGTCAGGCTGTGGAGAGTCTCTTCCAGGAGGCGAGCGGCTTTCGCTTTCGCT 1301
Db |||||
Qy 1261 TGCTTACAGCGCGACGCGGCTTCAAGTCTTGGGGCTGACGCTGCGCTGCGGCTCGTC 1320
Db |||||
Qy 1302 TGCTTACAGCGCGACGCGGCTTCAAGTCTTGGGGCTGACGCTGCGCTGCGGCTCGTC 1361
Db |||||
Qy 1321 GTCTTACGCTGTCTATGCGCATTTATGTCGCGCATCTTTCGCTGCTTTCGCTTTCGCT 1380
Db |||||
Qy 1362 GTCTTACGCTGTCTATGCGCATTTATGTCGCGCATCTTTCGCTGCTTTCGCTTTCGCT 1421
Db |||||
Qy 1381 GGAGCGCTACGCGGCGCGGCTTCTTCTGCTGCGGCTTTCGCTTTCGCTTTCGCT 1440
Db |||||
Qy 1422 GGAGCGCTACGCGGCGCGGCTTCTTCTGCTGCGGCTTTCGCTTTCGCTTTCGCT 1481
Db |||||
Qy 1441 CTCTGGCGCAAGCTGTCTGTCGACCAAGTCTTCTTTCGAGCTGCGCATCTTTCGCTTTCGCT 1500
Db |||||
Qy 1482 CTCTGGCGCAAGCTGTCTGTCGACCAAGTCTTCTTTCGAGCTGCGCATCTTTCGCTTTCGCT 1541
Db |||||
Qy 1501 GGCATCTGACGCTGTCGCGCTTCTGTCATCTCCCTGAGGGCTTCTATGAGGCTTACCGA 1560
Db |||||
Qy 1542 GGCATCTGACGCTGTCGCGCTTCTGTCATCTCCCTGAGGGCTTCTATGAGGCTTACCGA 1601
Db |||||
Qy 1561 ACCAAGCGGAGGACTAG 1578
Db |||||
Qy 1602 ACCAAGCGGAGGACTAG 1619
Db |||||

RESULT 12

AAS06087/c

ID AAS06087 standard; DNA; 1723 BP.

XX AAS06087;

XX AC

XX AC

XX 12-SEP-2001 (first entry)

XX Angiotensin converting enzyme (ACEV) splice variant DNA #87.

XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C; ds;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonarctoidotic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.

Qy 961 ATCTTCTGCTTCTGCTGGAGGCAATATGACAGAGCCAGAGTCTCCACTGCATGATG 1020
Db 663 ATCTTCTGCTTCTGCTGGAGGCAATATGACAGAGCCAGAGTCTCCACTGCATGATG 604
Qy 1021 AACTGGACGACATCGAGCGCTGCTGCTCAAGGGCCCTCTTGGCGCTGCTGCGCTACCTC 1080
Db 603 AACTGGACGACATCGAGCGCTGCTGCTCAAGGGCCCTCTTGGCGCTGCTGCGCTACCTC 544
Qy 1081 ACCTGGCCGACGAGACCAAGGAGTCTACCGATTAACCTGCGCGCTCCATCCGGGCC 1140
Db 543 ACCTGGCCGACGAGACCAAGGAGTCTACCGATTAACCTGCGCGCTCCATCCGGGCC 484
Qy 1141 GTGCTCAACATCTTTCTGTGGCCCAAGGCGCTGTGTCTATCTCTGCGCATCTTTGGCC 1200
Db 483 GTGCTCAACATCTTTCTGTGGCCCAAGGCGCTGTGTCTATCTCTGCGCATCTTTGGCC 424
Qy 1201 GCTGTCAGGTGCTGGAGAAGTCTCTTCCAGGAAGGCGCGCTTTTCCCGGCC 1260
Db 423 GCTGTCAGGTGCTGGAGAAGTCTCTTCCAGGAAGGCGCGCTTTTCCCGGCC 364
Qy 1261 TGCTACAGCGCGAGCGCGCTGAAGTCTCTGGGGCTGAGCTGCGCTGCGGCTGCTGTC 1320
Db 363 TGCTACAGCGCGAGCGCGCTGAAGTCTCTGGGGCTGAGCTGCGCTGCGGCTGCTGTC 304
Qy 1321 GTCTTACGCTGCTCATGGCCATTATGTGCGCATTCTGCGCTGCTCATGGGCTCACC 1380
Db 303 GTCTTACGCTGCTCATGGCCATTATGTGCGCATTCTGCGCTGCTCATGGGCTCACC 244
Qy 1381 GGCAGCTCACGGCGCGCGCTCTGTTTCTGCTGCCAGCGCTCTTTCACCTGCGCGTG 1440
Db 243 GGCAGCTCACGGCGCGCGCTCTGTTTCTGCTGCCAGCGCTCTTTCACCTGCGCGTG 184
Qy 1441 CTCTGGCGGAGCTGCTGTGGCACCAGTCTTCTTGACGTGCGCATCTTCTGTCATGGC 1500
Db 183 CTCTGGCGGAGCTGCTGTGGCACCAGTCTTCTTGACGTGCGCATCTTCTGTCATGGC 124
Qy 1501 GGCATCTGACGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
Db 123 GGCATCTGACGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 64
Qy 1561 ACCAAGCGGAGGACTAG 1578
Db 63 ACCAAGCGGAGGACTAG 46

RESULT 10
ABK92054/c

ID ABK92054 standard; DNA; 1734 BP.

XX AC ABK92054;

XX AC ABK92054;

XX DT 14-AUG-2002 (first entry)

XX DE DNA encoding novel GABA transporter-like receptor protein #6.
XX KW Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;
XX KW behavioral disorder; valve disease; endocrine disorder; heart disorder;
XX KW blood disorder; anxiety disorder; brain disease; inflammatory disorder;
XX KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX KW immune disorder; haematopoietic disorder; dyslipidaemia;
XX KW metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;
XX KW gene therapy; transgenic animal; human; gene; ds.

XX OS Homo sapiens.

XX PN WO200229058-A2.

XX PD 11-APR-2002.

XX PF 05-OCT-2001; 2001WO-US031248.

XX PR 05-OCT-2000; 2000US-0238323P.

XX PR 05-OCT-2000; 2000US-0238325P.

PR 06-OCT-2000; 2000US-0238372P.
PR 06-OCT-2000; 2000US-0238373P.
PR 06-OCT-2000; 2000US-0238379P.
PR 06-OCT-2000; 2000US-0238382P.
PR 06-OCT-2000; 2000US-0238383P.
PR 06-OCT-2000; 2000US-0238384P.
PR 06-OCT-2000; 2000US-0238397P.
PR 06-OCT-2000; 2000US-0238400P.
PR 06-OCT-2000; 2000US-0238401P.
PR 06-OCT-2000; 2000US-0238402P.
PR 14-MAR-2001; 2001US-0275892P.
PR 08-JUN-2001; 2001US-0296860P.
XX (CURA-) CURAGEN CORP.
XX PA
XX PI Shimkets RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;
PI Rastelli L, Malyankar UM, Grosse WM, Alsobrook JP, Lepley DM;
PI Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;
PI Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;
XX
DR WPI; 2002-444103/47.
DR P-PSDB; ABG61787.
XX
PT Novel isolated polypeptide, designated NOVX, useful for treating or
PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT metabolic, neurodegenerative, immune and hematopoietic disorders.
XX
PS Claim 9; Page 64; 316pp; English.
XX
CC The invention describes an isolated polypeptide (I), useful in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease. (I), the polynucleotide encoding it (II) and an antibody
CC (III) to (I) are useful for treating or preventing cancer, metabolic
CC disorders, skin disorders, infectious disease, anorexia, behavioral
CC disorders, valve diseases, endocrine disorders, heart and blood
CC disorders, anxiety disorders, brain disorders, inflammatory disorders,
CC neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,
CC immune disorders, haematopoietic disorders, and the various
CC dyslipidaemias, metabolic disturbances associated with obesity, and the
CC metabolic syndrome X and wasting disorders associated with chronic
CC diseases. (I), (II) or (III) are useful in screening assays, detection
CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC predictive medicine (e.g., diagnostic assays and prognostic assays), and
CC in methods of treatment (e.g., therapeutic and prophylactic). (I) is
CC useful as immunogen to produce antibodies immunospecific for (I), to
CC screen for potential agonist and antagonist compounds, and as bait
CC protein in a two-hybrid or three-hybrid assay. (II) is useful in gene
CC therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)
CC gene, and to modulate activity of (I). A cell containing a vector
CC expressing (I) is useful for producing non-human transgenic animals. This
CC sequence encodes a novel human polypeptide described in the invention
XX
SQ Sequence 1734 BP; 357 A; 514 C; 580 G; 283 T; 0 U; 0 Other;
Query Match 99.8%; Score 1574.8; DB 6; Length 1734;
Best Local Similarity 99.9%; Pred. No. 6.9e-311; Mismatches 0; Gaps 0;
Matches 1576; Conservative 0; Indels 2; Indels 0; Gaps 0;
Qy 1 ATGGCCACCTTGTCTCGCAGCAGCTGTCCAAAGTGGCCAGCTGTCCAAAGTTC 60
Db 1623 ATGGCCACCTTGTCTCGCAGCAGCTGTCCAAAGTGGCCAGCTGTCCAAAGTTC 1564
Qy 61 CAGGCCAAGATGAGCGGCATGTTTCCAGGATGGTGTTCAGCGCGCCAGGATGGAG 120
Db 1563 CAGGCCAAGATGAGCGGCATGTTTCCAGGATGGTGTTCAGCGCGCCAGGATGGAG 1504
Qy 121 GCGGTGGCTTCGCGCATTTGCGACGACCTCGACTTTTGGACCGCCAGGATGGAG 180
Db 1503 GCGGTGGCTTCGCGCATTTGCGACGACCTCGACTTTTGGACCGCCAGGATGGAG 1444
Qy 181 GACATCTGAAAGCCGAGGGAGAGCCCTGCGGGGACGAGGGCGCTGAAGCGCCGCTGAG 240
Db 1443 GACATCTGAAAGCCGAGGGAGAGCCCTGCGGGGACGAGGGCGCTGAAGCGCCGCTGAG 1384

AC	ABK92053;
XX	
DT	14-AUG-2002 (first entry)
XX	
DE	DNA encoding novel GABA transporter-like receptor protein #5.
XX	
KW	Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;
KW	behavioral disorder; valve disease; endocrine disorder; heart disorder;
KW	blood disorder; anxiety disorder; brain disorder; inflammatory disorder;
KW	neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW	immune disorder; haematopoietic disorder; dyslipidaemia;
KW	metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;
KW	gene therapy; transgenic animal; human; gene; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200229058-A2.
XX	
PD	11-APR-2002.
XX	
PF	05-OCT-2001; 2001WO-US031248.
XX	
PR	05-OCT-2000; 2000US-0238323P.
PR	05-OCT-2000; 2000US-0238325P.
PR	06-OCT-2000; 2000US-0238372P.
PR	06-OCT-2000; 2000US-0238373P.
PR	06-OCT-2000; 2000US-0238379P.
PR	06-OCT-2000; 2000US-0238382P.
PR	06-OCT-2000; 2000US-0238383P.
PR	06-OCT-2000; 2000US-0238384P.
PR	06-OCT-2000; 2000US-0238397P.
PR	06-OCT-2000; 2000US-0238400P.
PR	06-OCT-2000; 2000US-0238401P.
PR	06-OCT-2000; 2000US-0238402P.
PR	14-MAR-2001; 2001US-0275892P.
PR	08-JUN-2001; 2001US-0298608P.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
PI	Shimkets RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;
PI	Rastelli L, Malyankar UM, Grosse WM, Alsobrook JP, Lepley DM;
PI	Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;
PI	Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;
XX	
DR	WPI; 2002-444103/47.
DR	P-PSDB; ABG61786.
XX	
PT	Novel isolated polypeptide, designated NOVX, useful for treating or
PT	preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT	metabolic, neurodegenerative, immune and hematopoietic disorders.
XX	
PS	Claim 9; Page 63-64; 316pp; English.
XX	
CC	The invention describes an isolated polypeptide (I), useful in the
CC	manufacture of a medicament for treating a syndrome associated with a
CC	human disease. (II), the polynucleotide encoding it (II) and an antibody
CC	(III) to (I) are useful for treating or preventing cancer, metabolic
CC	disorders, skin disorders, infectious disease, anorexia, behavioral
CC	disorders, valve diseases, endocrine disorders, heart and blood
CC	disorders, anxiety disorders, brain disorders, inflammatory disorders,
CC	neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,
CC	immune disorders, haematopoietic disorders, and the various
CC	dyslipidemias, metabolic disturbances associated with obesity, and the
CC	metabolic syndrome X and wasting disorders associated with chronic
CC	diseases. (I), (II) or (III) are useful in screening assays, detection
CC	assays (e.g., chromosomal mapping, tissue typing, forensic biology), .
CC	predictive medicine (e.g., diagnostic assays and prognostic assays); and
CC	in methods of treatment (e.g., therapeutic and prophylactic). (I) is
CC	useful as immunogen to produce antibodies immunospecific for (I), to
CC	screen for potential agonist and antagonist compounds, and as bait
CC	protein in a two-hybrid or three-hybrid assay. (II) is useful in gene
CC	therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)
CC	gene, and to modulate activity of (I). A cell containing a vector

CC	expressing (I) is useful for producing non-human transgenic animals. This
CC	sequence encodes a novel human polypeptide described in the invention
XX	
SQ	Sequence 1734 BP; 358 A; 515 C; 579 G; 282 T; 0 U; 0 Other;
XX	
Query Match	99.8%; Score 1574.8; DB 6; Length 1734;
Best Local Similarity	99.9%; Pred. No. 6.9e-311;
Matches 1576; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
XX	
QY	1 ATGGCCACCTTGTCTCCGCAGCAAGCTGTCCAACGTGGCCACGTCCTGTGTCCAAACAATGCC 60
Db	1623 ATGGCCACCTTGTCTCCGCAGCAAGCTGTCCAACGTGGCCACGTCCTGTGTCCAAACAATGCC 1564
XX	
QY	61 CAGGCCAAGATAGCGGCATGTTCCGACAGATGGGTTTTTAGCGGGCACCAGATAGGAG 120
Db	1563 CAGGCCAAGATAGCGGCATGTTCCGACAGATGGGTTTTTAGCGGGCACCAGATAGGAG 1504
XX	
QY	121 GCGGTGGGCTTCGGGCATTGCGACGACCTCGACTTTTGAGCACACGCCAGGGCCTGCACATG 180
Db	1503 GCGGTGGGCTTCGGGCATTGCGACGACCTCGACTTTTGAGCACACGCCAGGGCCTGCACATG 1444
XX	
QY	181 GACATCCTGAAGCCGAGGAGAGCCCTGCGGGGACGAGGCGCTGAAGCGCCGTCGAG 240
Db	1443 GACATCCTGAAGCCGAGGAGAGCCCTGCGGGGACGAGGCGCTGAAGCGCCGTCGAG 1384
XX	
QY	241 GGAGACATCCATTATACGAGGAGGAGCGGAGCTCTCTGCGCCCTCTCGGCTTCAAAGAC 300
Db	1383 GGAGACATCCATTATACGAGGAGGAGCGGAGCTCTCTGCGCCCTCTCGGCTTCAAAGAC 1324
XX	
QY	301 CAGGTGGGAGGTGGTGGCGAATTCGGGGGCCACACAGAAGCCCAAAAATCAGGCGTGGGAG 360
Db	1323 CAGGTGGGAGGTGGTGGCGAATTCGGGGGCCACACAGAAGCCCAAAAATCAGGCGTGGGAG 1264
XX	
QY	361 GCAGGCTGGAACGCTGACCAACGCCCATCCAGGGCATGTTGCTGTGGGCTTACCCCTACGCC 420
Db	1263 GCAGGCTGGAACGCTGACCAACGCCCATCCAGGGCATGTTGCTGTGGGCTTACCCCTACGCC 1204
XX	
QY	421 ATCTGTGACGCGGCTACCTGGGGTGTGTTCTCATCATCTTTCGCGCGCGCTGTGTGCTGC 480
Db	1203 ATCTGTGACGCGGCTACCTGGGGTGTGTTCTCATCATCTTTCGCGCGCGCTGTGTGCTGC 1144
XX	
QY	481 TACACGGCAGATCCTCATCGCTGCTGTACGAGGAGATGAAGACGGGAGGTGGTG 540
Db	1143 TACACGGCAGATCCTCATCGCTGCTGTACGAGGAGATGAAGACGGGAGGTGGTG 1084
XX	
QY	541 CGCGTGGGAGCTCGTAGTGCCATAGCCAAAGCTGCTCGCGCCCGCGCTTCCCACAG 600
Db	1083 CGCGTGGGAGCTCGTAGTGCCATAGCCAAAGCTGCTCGCGCCCGCGCTTCCCACAG 1024
XX	
QY	601 CTGGGCGGCCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGATGACGTGATCCTG 660
Db	1023 CTGGGCGGCCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGATGACGTGATCCTG 964
XX	
QY	661 TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCCGCTGTCGAG 720
Db	963 TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCCGCTGTCGAG 904
XX	
QY	721 AAGTCTGTGTCATTATCGCACGGCCTGTGCTGCTTGGCGCTTCTTTAAGAACCTC 780
Db	903 AAGTCTGTGTCATTATCGCACGGCCTGTGCTGCTTGGCGCTTCTTTAAGAACCTC 844
XX	

PT metabolic, neurodegenerative, immune and hematopoietic disorders.
XX Claim 9; Page 63; 316pp; English.
XX
CC The invention describes an isolated polypeptide (I), useful in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease. (I), the polynucleotide encoding it (II) and an antibody
CC (III) to (I) are useful for treating or preventing cancer, metabolic
CC disorders, skin disorders, infectious diseases, anorexia, behavioral
CC disorders, valve diseases, endocrine disorders, heart and blood
CC neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,
CC immune disorders, haematopoietic disorders, and the various
CC dyslipidaemias, metabolic disturbances associated with obesity, and the
CC metabolic syndrome X and wasting disorders associated with chronic
CC diseases. (I), (II) or (III) are useful in screening assays, detection
CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC predictive medicine (e.g., diagnostic assays and prognostic assays), and
CC in methods of treatment (e.g., therapeutic and prophylactic). (I) is
CC useful as immunogen to produce antibodies immunospecific for (I), to
CC screen for potential agonist and antagonist compounds, and as bait
CC protein in a two-hybrid or three-hybrid assay. (II) is useful in gene
CC therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)
CC gene, and to modulate activity of (I). A cell containing a vector
CC expressing (I) is useful for producing non-human transgenic animals. This
CC sequence encodes a novel human polypeptide described in the invention
XX
SQ Sequence 1734 BP; 356 A; 515 C; 581 G; 282 T; 0 U; 0 Other;
Query Match 99.8%; Score 1574.8; DB 6; Length 1734;
Best Local Similarity 99.9%; Pred. No. 6.9e-311;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCCACCTTGCTCCGAGCAAGCTGTCCAACTGGGCCAGTCCGTGTCACCAAGTCC 60
Db 1623 ATGGCCACCTTGCTCCGAGCAAGCTGTCCAACTGGGCCAGTCCGTGTCACCAAGTCC 1564
Qy 61 CAGGCCAAGATAGCGGCATGTTCCGAGGATGGTTTTCAGCGGCCAGCGATGAGGAG 120
Db 1563 CAGGCCAAGATAGCGGCATGTTCCGAGGATGGTTTTCAGCGGCCAGCGATGAGGAG 1504
Qy 121 GCGGTGGGCTTTCGCGCATTCGACGACCTCGACTTTTGAGCACCCAGGCGCTGCAGATG 180
Db 1503 GCGGTGGGCTTTCGCGCATTCGACGACCTCGACTTTTGAGCACCCAGGCGCTGCAGATG 1444
Qy 181 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 240
Db 1443 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 1384
Qy 241 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTGCGGCTCCGGCTCCAGGAC 300
Db 1383 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTGCGGCTCCGGCTCCAGGAC 1324
Qy 301 CAGGTGGGAGGTGTGGCGAATTCGGGGGCCACGACAGGCCAAATATCAGCGGTGGGAG 360
Db 1323 CAGGTGGGAGGTGTGGCGAATTCGGGGGCCACGACAGGCCAAATATCAGCGGTGGGAG 1264
Qy 361 GCAGGCTGGAACTGACCAAGCCATCCAGGGCATGTTTCGTGCTGGGCCCTACCCCTACGCC 420
Db 1263 GCAGGCTGGAACTGACCAAGCCATCCAGGGCATGTTTCGTGCTGGGCCCTACCCCTACGCC 1204
Qy 421 ATCTGTCAGCGGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGGCTGTGTGCTGC 480
Db 1203 ATCTGTCAGCGGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGGCTGTGTGCTGC 1144
Qy 481 TACACCGGCAAGATCTCATCGCGTGTCTGTACGAGGAGATGAAGACGCGAGGTGGTG 540
Db 1143 TACACCGGCAAGATCTCATCGCGTGTCTGTACGAGGAGATGAAGACGCGAGGTGGTG 1084
Qy 541 CGCGTGGGAGCTCGTACGTGGCCATAGCCAAAGCCCTGTCGCGCCCGCGCTTCCCAACG 600
Db 1083 CGCGTGGGAGCTCGTACGTGGCCATAGCCAAAGCCCTGTCGCGCCCGCGCTTCCCAACG 1024

Qy 601 CTGGGCGGCGAGTGGTGAAGTAGCGAGATCATCGAGCTGGTGTAGTACGTGTCATCTG 660
Db 1023 CTGGGCGGCGAGTGGTGAAGTAGCGAGATCATCGAGCTGGTGTAGTACGTGTCATCTG 964
Qy 661 TACGTGGTGGTGGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCCGCTGTCGAG 720
Db 963 TACGTGGTGGTGGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCCGCTGTCGAG 904
Qy 721 AAGTCTGGTCCATTATTCGACGCGCGTGTGCTGCTGCGGCTTCCCTTAAGAACTC 780
Db 903 AAGTCTGGTCCATTATTCGACGCGCGTGTGCTGCTGCGGCTTCCCTTAAGAACTC 844
Qy 781 AAGGCCGTGTCGAAGTTCAGTCTGCTGCACTCTGGCCCACTTCGTCATCAATATCTG 840
Db 843 AAGGCCGTGTCGAAGTTCAGTCTGCTGCACTCTGGCCCACTTCGTCATCAATATCTG 784
Qy 841 GTCATAGCCTACTGTCTATCGCGGCGCGGCACTTGGGCTTGGGAGAAAGTCAAGTTCAC 900
Db 783 GTCATAGCCTACTGTCTATCGCGGCGCGGCACTTGGGCTTGGGAGAAAGTCAAGTTCAC 724
Qy 901 ATGACGTCAAGAAAGTTCCTCATCTCATTTGGCATCATCTGTTTCAAGTACAGTCTCAG 960
Db 723 ATGACGTCAAGAAAGTTCCTCATCTCATTTGGCATCATCTGTTTCAAGTACAGTCTCAG 664
Qy 961 ATCTTCTGCTTTCGCTGGAGGCGCAATATGACGAGCCGAGGAGTTCACATGATGATG 1020
Db 663 ATCTTCTGCTTTCGCTGGAGGCGCAATATGACGAGCCGAGGAGTTCACATGATGATG 604
Qy 1021 AACTGGACGCACATCGCAGCTCGTGTCTCAAGGGGCTCTTTCGCGCTCGTCTGCTACCTC 1080
Db 603 AACTGGACGCACATCGCAGCTCGTGTCTCAAGGGGCTCTTTCGCGCTCGTCTGCTACCTC 544
Qy 1081 ACTTGGGCCGACGAGACCAAGAGGTCATCAACGATAAATCTGCGCGGCTCCATCCGGGCC 1140
Db 543 ACTTGGGCCGACGAGACCAAGAGGTCATCAACGATAAATCTGCGCGGCTCCATCCGGGCC 484
Qy 1141 GTGCTCAACATCTTCTGTTGGGCAAGGCGCTGTGCTATCTCTGCGCATTTCTTTGTC 1200
Db 483 GTGCTCAACATCTTCTGTTGGGCAAGGCGCTGTGCTATCTCTGCGCATTTCTTTGTC 424
Qy 1201 GCTGTGAGGTGCTGGAGAGTCCCTTTCAGAGAAAGGACGCGGCTTTTTCGCGGCC 1260
Db 423 GCTGTGAGGTGCTGGAGAGTCCCTTTCAGAGAAAGGACGCGGCTTTTTCGCGGCC 364
Qy 1261 TGCTACAGCGCGACGCGGCGCTTGAAGTCTCTGGGGGTGACGCTGCGCTGCGGCTGCTC 1320
Db 363 TGCTACAGCGCGACGCGGCGCTTGAAGTCTCTGGGGGTGACGCTGCGCTGCGGCTGCTC 304
Qy 1321 GTCTTCAAGCTGCTCATGGCATTATGTGCGGCACTTTCGCGCTGCTCATGGGCTCACC 1380
Db 303 GTCTTCAAGCTGCTCATGGCATTATGTGCGGCACTTTCGCGCTGCTCATGGGCTCACC 244
Qy 1381 GGCAGCTTCAAGGCGCGCGGCTCTGTCTTCTGCTGCCAGGCTCTTTCACTGCGGCTG 1440
Db 243 GGCAGCTTCAAGGCGCGCGGCTCTGTCTTCTGCTGCCAGGCTCTTTCACTGCGGCTG 184
Qy 1441 CTCTGCGCAAGCTGCTGGGACCAAGTCTTCTTCAAGTTCGCCATCTTTTCGTATCGGC 1500
Db 183 CTCTGCGCAAGCTGCTGGGACCAAGTCTTCTTCAAGTTCGCCATCTTTTCGTATCGGC 124
Qy 1501 GGCATCTGAGGCTGCTGCGGCTTCTGCTGCACTCCCTCGAGGGGCTCATCGAGGCTACCGA 1560
Db 123 GGCATCTGAGGCTGCTGCGGCTTCTGCTGCACTCCCTCGAGGGGCTCATCGAGGCTACCGA 64
Qy 1561 ACCAACGCGGAGGACTAG 1578
Db 63 ACCAACGCGGAGGACTAG 46

RESULT 9
ABK92053/C
ID ABK92053 standard; DNA; 1734 BP.
XX

DT	09-APR-2002 (first entry)		CC	invention. This gene maps to human chromosome 20q12-20q13
XX			XX	
DE	Human vesicular GABA transporter (VGAT) gene sequence.		SQ	Sequence 1800 BP; 290 A; 613 C; 525 G; 372 T; 0 U; 0 Other;
XX				
KW	Human; vesicular GABA transporter; VGAT; gamma aminobutyric acid; GABA;			Query Match 99.9%; Score 1576.4; DB 5; Length 1800;
KW	neurotrophic; neuroprotective; antitumor; anticonvulsant; analgesic;			Best Local Similarity 99.9%; Pred. No. 3.3e-311;
KW	antiinflammatory; neuroleptic; chromosome localisation; antidepressant;			Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
KW	screening assay; agonist; antagonist; schizophrenia; epilepsy;		QY	1 ATGGCCACTTGTCTCCGACGAGCTGTCACCAAGTGGCCAGCTCCGTCGTCCAAACAGTCC 60
KW	depression; learning disorder; pain; cognitive disorder;		DB	100 ATGGCCACTTGTCTCCGACGAGCTGTCACCAAGTGGCCAGCTCCGTCGTCCAAACAGTCC 159
KW	neurodegenerative disease; multiple sclerosis; dementia;		QY	61 CAGGCCAAGATGAGCGGATGTTCCGCCAGGATGGTTTTCAGGGCGCCACGATGAGGAG 120
KW	Alzheimer's disease; Parkinson's disease; Crohn's disease;		DB	160 CAGGCCAAGATGAGCGGATGTTCCGCCAGGATGGTTTTCAGGGCGCCACGATGAGGAG 219
KW	anxiety disorder; sleeping disorder; alcoholism; muscular disorder;		QY	121 CGGTGGGCTTCGGCGCATTTGCGACCTGACCTTTAGACACCGCCAGGGCTTCAGATG 180
KW	tremor; headache; migraine; immunogen; membrane bound receptor;		DB	220 CGGTGGGCTTCGGCGCATTTGCGACCTGACCTTTAGACACCGCCAGGGCTTCAGATG 279
KW	soluble receptor; hybridisation probe; primer; diagnostic assay;			
XX	tissue expression; transgenic animal; chromosome 20q12-20q13; ds.			
OS	Homo sapiens.			
XX				
XX		Location/Qualifiers		
PH	Key	100..1677		
FT	CDS	/*tag= a		
FT		/product= "Human VGAT"		
FT				
XX	WO200173015-A1.			
XX	04-OCT-2001.			
XX	23-MAR-2001; 2001WO-BP003350.			
XX	27-MAR-2000; 2000EP-00106581.			
XX	(MERE) MERCK PATENT GMBH.			
PA	Rippmann F, Duecker K;			
PI	WPI: 2001-616495/71.			
XX	P-PSDB; AAU09936.			
DR				
DR				
XX	New polypeptide for diagnosing, treating neurodegenerative diseases,			
PT	epilepsy, pain, sleeping, anxiety disorder and identifying modulators			
PT	comprises the human vesicular gamma aminobutyric acid transporter			
PT	polypeptide.			
XX				
PS	Claim 5; Page 31-33; 39pp; English.			
XX				
CC	The invention relates to a new polypeptide comprising a fully defined			
CC	sequence of 525 amino acids as given in the specification, or a			
CC	polypeptide comprising a sequence 95% homologous to the fully defined			
CC	sequence of the human vesicular gamma aminobutyric acid transporter,			
CC	(VGAT). The invention also relates to fragments or variants of VGAT,			
CC	encoded by a polynucleotide comprising a fully defined sequence of 1800			
CC	base pairs as given in the specification. The VGAT polypeptide is useful			
CC	in screening assays to identify compounds that stimulate or inhibit the			
CC	function or level of the polypeptide. VGAT polypeptides, polynucleotides			
CC	and the agonists and antagonists identified by the above method are			
CC	useful for treating schizophrenia, epilepsy, depression, learning			
CC	disorders, cognitive disorders, neurodegenerative diseases, multiple			
CC	sclerosis, dementia, Alzheimer's, Parkinson's disease, Crohn's disease,			
CC	ulcerative colitis, dyspepsia, irritable bowel syndrome, hyperactivity,			
CC	anxiety disorder, sleeping disorder, alcoholism, muscular disorders e.g.			
CC	tremor, pain, headache and migraine. The VGAT polypeptides and			
CC	polynucleotides are also useful as vaccines and as immunogens to produce			
CC	antibodies which are useful for treating diseases. VGAT polypeptides can			
CC	be used to identify membrane bound or soluble receptors and as			
CC	hybridisation probes for cDNA and genomic DNA, as primers for nucleic			
CC	acid amplification reactions to isolate full-length cDNAs and genomic			
CC	clones encoding VGAT polypeptides, in diagnostic assays by detecting			
CC	mutations in the associated gene, for chromosome localisation studies,			
CC	tissue expression studies and for producing transgenic animals useful in			
CC	drug discovery and target validation. The present sequence encodes the			
CC	human vesicular gamma aminobutyric acid transporter (VGAT) protein of the			

PS Claim 9; Page 60-61; 316pp; English.

xx The invention describes an isolated polypeptide (I), useful in the manufacture of a medicament for treating a syndrome associated with a human disease. (I), the polynucleotide encoding it (II) and an antibody (III) to (I) are useful for treating or preventing cancer, metabolic disorders, skin disorders, infectious disease, anorexia, behavioral disorders, valve diseases, endocrine disorders, heart and blood disorders, anxiety disorders, brain disorders, inflammatory disorders, neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder, immune disorders, haematopoietic disorders, and the various dyslipidaemias, metabolic disturbances associated with obesity, and the metabolic syndrome X and wasting disorders associated with chronic diseases. (I), (II) or (III) are useful in screening assays, detection assays (e.g., chromosomal mapping, tissue typing, forensic biology), and predictive medicine (e.g., diagnostic assays and prognostic assays), and in methods of treatment (e.g., therapeutic and prophylactic). (I) is useful as immunogen to produce antibodies immunospecific for (I), to screen for potential agonist and antagonist compounds, and as bait protein in a two-hybrid or three-hybrid assay. (II) is useful in gene therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I) gene, and to modulate activity of (I). A cell containing a vector expressing (I) is useful for producing non-human transgenic animals. This sequence encodes a novel human polypeptide described in the invention

xx

Sequence 1763 BP; 286 A; 590 C; 527 G; 360 T; 0 U; 0 Other;

Query Match 99.9%; Score 1576.4; DB 6; Length 1763;
Best Local Similarity 99.9%; Pred. No. 3.3e-311;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCCACCTTGTCCGAGCAAGCTGTCCAAAGTGGCCAGTCCGCTGCCAACAGTCC 60
Db 141 ATGCCACCTTGTCCGAGCAAGCTGTCCAAAGTGGCCAGTCCGCTGCCAACAGTCC 200

Qy 61 CAGGCCAAGATGAGCGGATGTTCCGAGGATGGGTTTTCAGGGGCGCCAGGATGAGGAG 120
Db 201 CAGGCCAAGATGAGCGGATGTTCCGAGGATGGGTTTTCAGGGGCGCCAGGATGAGGAG 260

Qy 121 CGGTGGGCTTCGGCGATTCGAGACCTCGACTTTGAGACACGCCAGGGCTGCGAGATG 180
Db 261 CGGTGGGCTTCGGCGATTCGAGACCTCGACTTTGAGACACGCCAGGGCTGCGAGATG 320

Qy 181 GACATCTGAAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 240
Db 321 GACATCTGAAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 380

Qy 241 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTGCGCCCTCCGGCTCCAAGGAC 300
Db 381 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTGCGCCCTCCGGCTCCAAGGAC 440

Qy 301 CAGGTGGAGGTGTGGCGAATTCGGGGGCGACGAAAGCCCAAAATCAAGGCGTGGAG 360
Db 441 CAGGTGGAGGTGTGGCGAATTCGGGGGCGACGAAAGCCCAAAATCAAGGCGTGGAG 500

Qy 361 GCAGCTGGAACGTGACCAAGCCATCCAGGCGATGTTCTGCTGGGCTACCTACCGCC 420
Db 501 GCAGCTGGAACGTGACCAAGCCATCCAGGCGATGTTCTGCTGGGCTACCTACCGCC 560

Qy 421 ATCTGACCGCGGCTACCTGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGTGTCG 480
Db 561 ATCTGACCGCGGCTACCTGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGTGTCG 620

Qy 481 TACACCGGCAAGATCTCATCGCTGCTGTGAGAGAGATGAAGACGCGAGGTGGT 540
Db 621 TACACCGGCAAGATCTCATCGCTGCTGTGAGAGAGATGAAGACGCGAGGTGGT 680

Qy 541 CGGTGGGAGCTGTAAGTGGCCATAGCCAGCTGCTGCGCCCGCGCTTCCCAAG 600
Db 681 CGGTGGGAGCTGTAAGTGGCCATAGCCAGCTGCTGCGCCCGCGCTTCCCAAG 740

Qy 601 CTGGGCGGCGAGTGGTGAACGTAGCGCAGATCATTCGAGCTGGTGTGATGAGCTGATCCTG 660

Db 741 CTGGGCGGCGAGTGGTGAACCTAGCGAGATCATTCGAGCTGGTGTGATGACGTGATCCTG 800

Qy 741 CTGGGCGGCGAGTGGTGAACCTAGCGAGATCATTCGAGCTGGTGTGATGACGTGATCCTG 800

Db 761 TACGTGGTGGTGAAGTGGCAACCTCATGTACAAGCTTCCCGGGCTGCGGTGCGAG 720

Qy 761 TACGTGGTGGTGAAGTGGCAACCTCATGTACAAGCTTCCCGGGCTGCGGTGCGAG 720

Db 801 TACGTGGTGGTGAAGTGGCAACCTCATGTACAAGCTTCCCGGGCTGCGGTGCGAG 860

Qy 801 TACGTGGTGGTGAAGTGGCAACCTCATGTACAAGCTTCCCGGGCTGCGGTGCGAG 860

Db 721 AAGTCCCTGGTTCATTATTCGACCGCGCTGTGCTGCTTGGCGCTTCCCTTAAGAACCTC 780

Qy 721 AAGTCCCTGGTTCATTATTCGACCGCGCTGTGCTGCTTGGCGCTTCCCTTAAGAACCTC 780

Db 861 AAGTCCCTGGTTCATTATTCGACCGCGCTGTGCTGCTTGGCGCTTCCCTTAAGAACCTC 920

Qy 861 AAGTCCCTGGTTCATTATTCGACCGCGCTGTGCTGCTTGGCGCTTCCCTTAAGAACCTC 920

Db 781 AAGCCCGTGTCCAAAGTTTCAGTCTGTGCACTCTGGCCCACTTTCGTCATCAATATCCTG 840

Qy 781 AAGCCCGTGTCCAAAGTTTCAGTCTGTGCACTCTGGCCCACTTTCGTCATCAATATCCTG 840

Db 921 AAGCCCGTGTCCAAAGTTTCAGTCTGTGCACTCTGGCCCACTTTCGTCATCAATATCCTG 980

Qy 921 AAGCCCGTGTCCAAAGTTTCAGTCTGTGCACTCTGGCCCACTTTCGTCATCAATATCCTG 980

Db 841 GTCATACCTTACTGTCTATTCGCGGCGCGGACTGGGCTTGGGAGAGGTCAAGTCTTAC 900

Qy 841 GTCATACCTTACTGTCTATTCGCGGCGCGGACTGGGCTTGGGAGAGGTCAAGTCTTAC 900

Db 981 GTCATACCTTACTGTCTATTCGCGGCGCGGACTGGGCTTGGGAGAGGTCAAGTCTTAC 1040

Qy 981 GTCATACCTTACTGTCTATTCGCGGCGCGGACTGGGCTTGGGAGAGGTCAAGTCTTAC 1040

Db 901 ATGACGCTCAAGAAAGTTTCCCATCTCCATTTGGGATCATCTGTTTTCAGCTACAGTCTCAG 960

Qy 901 ATGACGCTCAAGAAAGTTTCCCATCTCCATTTGGGATCATCTGTTTTCAGCTACAGTCTCAG 960

Db 1041 ATGACGCTCAAGAAAGTTTCCCATCTCCATTTGGGATCATCTGTTTTCAGCTACAGTCTCAG 1100

Qy 1041 ATGACGCTCAAGAAAGTTTCCCATCTCCATTTGGGATCATCTGTTTTCAGCTACAGTCTCAG 1100

Db 961 ATCTTCTGCTTTCGCTGGAGGCAATATGACGAGCCAGCCAGAGTTCCTTCCATTCGATGATG 1020

Qy 961 ATCTTCTGCTTTCGCTGGAGGCAATATGACGAGCCAGCCAGAGTTCCTTCCATTCGATGATG 1020

Db 1101 ATCTTCTGCTTTCGCTGGAGGCAATATGACGAGCCAGCCAGAGTTCCTTCCATTCGATGATG 1160

Qy 1101 ATCTTCTGCTTTCGCTGGAGGCAATATGACGAGCCAGCCAGAGTTCCTTCCATTCGATGATG 1160

Db 1021 AACTGAGCGCATTCGAGGCTGTGCTCAAGGGCTCTTTCGGCTGCTGCTGCTTTCCTC 1080

Qy 1021 AACTGAGCGCATTCGAGGCTGTGCTCAAGGGCTCTTTCGGCTGCTGCTGCTTTCCTC 1080

Db 1161 AACTGAGCGCATTCGAGGCTGTGCTCAAGGGCTCTTTCGGCTGCTGCTGCTTTCCTC 1220

Qy 1161 AACTGAGCGCATTCGAGGCTGTGCTCAAGGGCTCTTTCGGCTGCTGCTGCTTTCCTC 1220

Db 1081 ACCTGGGCGCACGAGACCAAGAGGTTCATCAGGATTAACCTGCCCGGCTTCATCCCGGCC 1140

Qy 1081 ACCTGGGCGCACGAGACCAAGAGGTTCATCAGGATTAACCTGCCCGGCTTCATCCCGGCC 1140

Db 1221 ACCTGGGCGCACGAGACCAAGAGGTTCATCAGGATTAACCTGCCCGGCTTCATCCCGGCC 1280

Qy 1221 ACCTGGGCGCACGAGACCAAGAGGTTCATCAGGATTAACCTGCCCGGCTTCATCCCGGCC 1280

Db 1141 GTGGTCAACATCTTTCGTGGGCGCTGTGCTTTCATCTTCCTTTCCTTTCCTTTCCTTTC 1200

Qy 1141 GTGGTCAACATCTTTCGTGGGCGCTGTGCTTTCATCTTCCTTTCCTTTCCTTTCCTTTC 1200

Db 1281 GTGGTCAACATCTTTCGTGGGCGCTGTGCTTTCATCTTCCTTTCCTTTCCTTTCCTTTC 1340

Qy 1281 GTGGTCAACATCTTTCGTGGGCGCTGTGCTTTCATCTTCCTTTCCTTTCCTTTCCTTTC 1340

Db 1201 GCTGTGAGGTGCTGAGAAAGTGTCTTTCAGGAAAGGAGCGCGGCTTTCCTTTCCTTTCCT 1260

Qy 1201 GCTGTGAGGTGCTGAGAAAGTGTCTTTCAGGAAAGGAGCGCGGCTTTCCTTTCCTTTCCT 1260

Db 1341 GCTGTGAGGTGCTGAGAAAGTGTCTTTCAGGAAAGGAGCGCGGCTTTCCTTTCCTTTCCT 1400

Qy 1341 GCTGTGAGGTGCTGAGAAAGTGTCTTTCAGGAAAGGAGCGCGGCTTTCCTTTCCTTTCCT 1400

Db 1261 TGCTACAGCGGCGACGCGGCTGAAAGTCTTGGGGCTGACGCTGCGCTGCGGCTGCTC 1320

Qy 1261 TGCTACAGCGGCGACGCGGCTGAAAGTCTTGGGGCTGACGCTGCGCTGCGGCTGCTC 1320

Db 1401 TGCTACAGCGGCGACGCGGCTGAAAGTCTTGGGGCTGACGCTGCGCTGCGGCTGCTC 1460

Qy 1401 TGCTACAGCGGCGACGCGGCTGAAAGTCTTGGGGCTGACGCTGCGCTGCGGCTGCTC 1460

Db 1321 GTCTTTCAGCTGCTCATGGCCATTTATGTGCGCAGCTTTCGCGCTGCTTCATGGGCTCACC 1380

Qy 1321 GTCTTTCAGCTGCTCATGGCCATTTATGTGCGCAGCTTTCGCGCTGCTTCATGGGCTCACC 1380

Db 1461 GTCTTTCAGCTGCTCATGGCCATTTATGTGCGCAGCTTTCGCGCTGCTTCATGGGCTCACC 1520

Qy 1461 GTCTTTCAGCTGCTCATGGCCATTTATGTGCGCAGCTTTCGCGCTGCTTCATGGGCTCACC 1520

Db 1381 GGAGCGCTCAAGCGGCGCGGCTTCTGTTTTCGCTGCGCAGCTTTCCTTTCACCTGCGCTG 1440

Qy 1381 GGAGCGCTCAAGCGGCGCGGCTTCTGTTTTCGCTGCGCAGCTTTCCTTTCACCTGCGCTG 1440

Db 1521 GGAGCGCTCAAGCGGCGCGGCTTCTGTTTTCGCTGCGCAGCTTTCCTTTCACCTGCGCTG 1580

Qy 1521 GGAGCGCTCAAGCGGCGCGGCTTCTGTTTTCGCTGCGCAGCTTTCCTTTCACCTGCGCTG 1580

Db 1441 CTCTGCGGCAAGCTGTGTGGGCAAGTCTTCTTCGAGCTGCGCATCTTTCGTCATCGGC 1500

Qy 1441 CTCTGCGGCAAGCTGTGTGGGCAAGTCTTCTTCGAGCTGCGCATCTTTCGTCATCGGC 1500

Db 1581 CTCTGCGGCAAGCTGTGTGGGCAAGTCTTCTTCGAGCTGCGCATCTTTCGTCATCGGC 1640

Qy 1581 CTCTGCGGCAAGCTGTGTGGGCAAGTCTTCTTCGAGCTGCGCATCTTTCGTCATCGGC 1640

Db 1501 GGATCTGCAAGCTGTTCGGCTTTCGTCATCTTTCGAGGCGCTTCATTCGAGGCTTACCGA 1560

Qy 1501 GGATCTGCAAGCTGTTCGGCTTTCGTCATCTTTCGAGGCGCTTCATTCGAGGCTTACCGA 1560

Db 1641 GGATCTGCAAGCTGTTCGGCTTTCGTCATCTTTCGAGGCGCTTCATTCGAGGCTTACCGA 1700

Qy 1641 GGATCTGCAAGCTGTTCGGCTTTCGTCATCTTTCGAGGCGCTTCATTCGAGGCTTACCGA 1700

Db 1561 ACCAAGCGGAGGACTAG 1578

Qy 1561 ACCAAGCGGAGGACTAG 1578

Db 1701 ACCAAGCGGAGGACTAG 1718

Qy 1701 ACCAAGCGGAGGACTAG 1718

RESULT 6
AAS15704
ID AAS15704 standard; DNA; 1800 BP.
xx
AC AAS15704;
XX

1383 GGAGATCCATTCATCAGCGAGCGAGAGCTCTCTGCGCGCTCCAGGAC 1324
301 CAGGTGGAGGTGGTGGCAATTCGGGGGCGACAGCAAGCCCAAAATCAAGCGGTGGAG 360
1323 CAGGTGGAGGTGGTGGCAATTCGGGGGCGACAGCAAGCCCAAAATCAAGCGGTGGAG 1264
361 CGAGGCTGGAAAGTGAACCAAGCCATCAGAGGCAATGTTCTGCTGGGCTACCCCTACGCC 420
1263 GCAGGCTGGAAAGTGAACCAAGCCATCAGAGGCAATGTTCTGCTGGGCTACCCCTACGCC 1204
421 ATCTGCAAGCGGGCTACTCTGGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGTGCTGC 480
1203 ATCTGCAAGCGGGCTACTCTGGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGTGCTGC 1144
481 TACACCGCAAGATCCTCATCGCTGCTGTACAGAGGAGATGAAGACGGGAGGTGTG 540
1143 TACACCGCAAGATCCTCATCGCTGCTGTACAGAGGAGATGAAGACGGGAGGTGTG 1084
541 CGCGTGGGACTCGTACGTGGCCATAGCCAAAGCTGCTGCGGCCCGCGCTTCCCAAG 600
1083 CGCGTGGGACTCGTACGTGGCCATAGCCAAAGCTGCTGCGGCCCGCGCTTCCCAAG 1024
601 CTGGGCGGCGAGTGGTGAACGTAGCGCAATCATCGAGCTGGTGAATGATGATCTCTG 660
1023 CTGGGCGGCGAGTGGTGAACGTAGCGCAATCATCGAGCTGGTGAATGATGATCTCTG 964
661 TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCGGTGCGAG 720
963 TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCGGTGCGAG 904
721 AAGTCTGCTGCTATTCGCCACGCGTGTGCTGCTTGGCGCTTCCCTTAAGAACCTC 780
903 AAGTCTGCTGCTATTCGCCACGCGTGTGCTGCTTGGCGCTTCCCTTAAGAACCTC 844
781 AAGCGCGTGCAGTTCAGTCTGTGTGCACTCTGGGCCCACTTGTGTATCAATATCTG 840
843 AAGCGCGTGCAGTTCAGTCTGTGTGCACTCTGGGCCCACTTGTGTATCAATATCTG 784
841 GTCATAGCTACTGTCTATCGCGGGCGCGAGCTGGGCTGGAGAGGTCAAGTTCTAC 900
783 GTCATAGCTACTGTCTATCGCGGGCGCGAGCTGGGCTGGAGAGGTCAAGTTCTAC 724
901 ATCGAGCTCAAGAAGTTCCTCCATCTCATTTGGCATCATCGTGTTCAGCTACAGTCTCAG 960
723 ATCGAGCTCAAGAAGTTCCTCCATCTCATTTGGCATCATCGTGTTCAGCTACAGTCTCAG 664
961 ATCTTCTGCTGCTGGAGGCAATATGCAAGAGCCAGCGAGTTCACATGCAATGATG 1020
663 ATCTTCTGCTGCTGGAGGCAATATGCAAGAGCCAGCGAGTTCACATGCAATGATG 604
1021 AACTGGACGACATCGACGCTGCGTGTCTAAGGGCTCTTTCGCGCTGCTGCGCTACCTC 1080
603 AACTGGACGACATCGACGCTGCGTGTCTAAGGGCTCTTTCGCGCTGCTGCGCTACCTC 544
1081 AACTGGGCGAGAGCAAGAGGTTCATCAGGATAAAGCTGCGGGCTCCATCGCGGCC 1140
543 AACTGGGCGAGAGCAAGAGGTTCATCAGGATAAAGCTGCGGGCTCCATCGCGGCC 484
1141 GTGGTCAACATCTTCTGGTGCCAAAGGCGCTGTTGTCTATCTCTGCGCATCTTTGCG 1200
483 GTGGTCAACATCTTCTGGTGCCAAAGGCGCTGTTGTCTATCTCTGCGCATCTTTGCG 424
1201 GCTGTGAGGTGCTGAGAGTGTCTTTCAGGAGGAGCGCGGCTTTTTCGCGGCC 1260
423 GCTGTGAGGTGCTGAGAGTGTCTTTCAGGAGGAGCGCGGCTTTTTCGCGGCC 364
1261 TGCTACAGCGGAGCGGCGCTTGAAGTCTGCGGGCTGACGCTGCGGTGCGGTCTGCTC 1320
363 TGCTACAGCGGAGCGGCGCTTGAAGTCTGCGGGCTGACGCTGCGGTGCGGTCTGCTC 304
1321 GTCTTACGCTGCTCATGGCCATTTATGTGCGGCACTTTCGCGTGTCTATGGGCTTACC 1380
303 GTCTTACGCTGCTCATGGCCATTTATGTGCGGCACTTTCGCGTGTCTATGGGCTTACC 244

Qy 1381 GGACGCTCAGGGCGCGGCTCTGTTTCTTGTGCTGCCAGCCTTTTCACTGCGCCTG 1440
Db 243 GGACGCTCAGGGCGCGGCTCTGTTTCTTGTGCTGCCAGCCTTTTCACTGCGCCTG 184
Qy 1441 CTCTGGGCGAAGCTGCTGTGGCACAAGTCTTCTTTCAGAGTGCACATCTTGTCTATCGC 1500
Db 183 CTCTGGGCGAAGCTGCTGTGGCACAAGTCTTCTTTCAGAGTGCACATCTTGTCTATCGC 124
Qy 1501 GGACATCTGACGCTGCTGGCTTGTGCACTCCCTCGAGGGCTTTCATCGAAGCCTTACCGA 1560
Db 123 GGACATCTGACGCTGCTGGCTTGTGCACTCCCTCGAGGGCTTTCATCGAAGCCTTACCGA 64
Qy 1561 ACCAACGCGAGGACTAG 1578
Db 63 ACCAACGCGAGGACTAG 46

RESULT 5
ABK92049

ID ABK92049 standard; DNA; 1763 BP.

XX ABK92049;

XX 14-AUG-2002 (first entry)

XX DNA encoding novel GABA transporter-like receptor protein #1.

XX Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;
XX behavioral disorder; valve disease; endocrine disorder; heart disorder;
XX blood disorder; anxiety disorder; brain disorder; inflammatory disorder;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX immune disorder; hematopoietic disorder; dyslipidaemia;
XX metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;
XX gene therapy; transgenic animal; human; gene; ds.

XX Homo sapiens.

XX WO200229058-A2.

XX 11-APR-2002.

XX 05-OCT-2001; 2001WO-US031248.

XX 05-OCT-2000; 2000US-0238323P.

XX 05-OCT-2000; 2000US-0238325P.

XX 06-OCT-2000; 2000US-0238372P.

XX 06-OCT-2000; 2000US-0238373P.

XX 06-OCT-2000; 2000US-0238379P.

XX 06-OCT-2000; 2000US-0238382P.

XX 06-OCT-2000; 2000US-0238383P.

XX 06-OCT-2000; 2000US-0238384P.

XX 06-OCT-2000; 2000US-0238397P.

XX 06-OCT-2000; 2000US-0238400P.

XX 06-OCT-2000; 2000US-0238401P.

XX 14-MAR-2001; 2001US-0275892P.

XX 08-JUN-2001; 2001US-0296860P.

XX (CURA-) CURAGEN CORP.

XX Shinketsu RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;

XX Rastelli L, Malvankar UM, Grosse WM, Alsobrook JP, Lepley DM;

XX Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;

XX Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;

XX WPI; 2002-444103/47.

XX P-PSDB; ABG61782.

XX Novel isolated polypeptide, designated NOVX, useful for treating or

XX preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and

XX metabolic, neurodegenerative, immune and hematopoietic disorders.

QY 1021 AACTGGACGACATCGCAGCTGCGTCTCAAGGGCTCTTTGGCGTCGCTGCTACCTC 1080
DB 1429 AACTGGACGACATCGCAGCTGCGTCTCAAGGGCTCTTTGGCGTCGCTGCTACCTC 1488
QY 1081 ACCTGGCCCGACGAGACCAAGGAGGTTCATCACGGATAACCTGCGCGCTCCATCCGCGCC 1140
DB 1489 ACCTGGCCCGACGAGACCAAGGAGGTTCATCACGGATAACCTGCGCGCTCCATCCGCGCC 1548
QY 1141 GTGGTCAACATCTTTCTGGTGGCAAGGCGCTGTGTCTTATCTCTGCGCATCTTTTGGC 1200
DB 1549 GTGGTCAACATCTTTCTGGTGGCAAGGCGCTGTGTCTTATCTCTGCGCATCTTTTGGC 1608
QY 1201 GCTGTCCAGGTGCTGGAGAGTCTCTTCCAGGAGGAGCGCGCTTTTTCGGGCC 1260
DB 1609 GCTGTCCAGGTGCTGGAGAGTCTCTTCCAGGAGGAGCGCGCTTTTTCGGGCC 1668
QY 1261 TGCTACAGCGCGACGCGCGCTCTGAAGTCTTGGGGCTGACGCTGCGCTGCGCGCTGTC 1320
DB 1669 TGCTACAGCGCGACGCGCGCTCTGAAGTCTTGGGGCTGACGCTGCGCTGCGCGCTGTC 1728
QY 1321 GTCTTACGCTGCTATGGCCATTTATGTGCGGACATTCGCGTGTCTATGGGCTTAC 1380
DB 1729 GTCTTACGCTGCTATGGCCATTTATGTGCGGACATTCGCGTGTCTATGGGCTTAC 1788
QY 1381 GGCAGCTCTACGGCGCGCGCTCTGTTCTTCTGTCGCCAGCCTCTTACCTGCGCGCTG 1440
DB 1789 GGCAGCTCTACGGCGCGCGCTCTGTTCTTCTGTCGCCAGCCTCTTACCTGCGCGCTG 1848
QY 1441 CTCTGGCGCAAGTGTCTGGCCACCAAGTCTTCTTCCAGCTGCGCATCTTTCGTCATCGGC 1500
DB 1849 CTCTGGCGCAAGTGTCTGGCCACCAAGTCTTCTTCCAGCTGCGCATCTTTCGTCATCGGC 1908
QY 1501 GGCATCTGCAGCTGTCGGCGCTTCGTGCACTTCCCTCGAGGCTCTATCGAAGCTTACCGA 1560
DB 1909 GGCATCTGCAGCTGTCGGCGCTTCGTGCACTTCCCTCGAGGCTCTATCGAAGCTTACCGA 1968
QY 1561 ACCAACGCGGAGACTAG 1578
DB 1969 ACCAACGCGGAGACTAG 1986
RESULT 4
ID ABK92051/C
XX ABK92051 standard; DNA; 1734 BP.
AC ABK92051;
XX
DT
XX
XX 14-AUG-2002 (first entry)
XX
DE DNA encoding novel GABA transporter-like receptor protein #3.
XX
KW Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;
KW behavioral disorder; valve disease; endocrine disorder; heart disorder;
KW blood disorder; anxiety disorder; brain disorder; inflammatory disorder;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia;
KW metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;
KW gene therapy; transgenic animal; human; gene; ds.
XX
OS Homo sapiens.
XX
XX WO200229058-A2.
XX
PD 11-APR-2002.
XX
XX 05-OCT-2001; 2001WO-US031248.
XX
XX 05-OCT-2000; 2000US-0238323P.
PR 05-OCT-2000; 2000US-0238325P.
PR 06-OCT-2000; 2000US-0238372P.
PR 06-OCT-2000; 2000US-0238373P.
PR 06-OCT-2000; 2000US-0238379P.

PR 06-OCT-2000; 2000US-0238382P.
PR 06-OCT-2000; 2000US-0238383P.
PR 06-OCT-2000; 2000US-0238384P.
PR 06-OCT-2000; 2000US-0238397P.
PR 06-OCT-2000; 2000US-0238400P.
PR 06-OCT-2000; 2000US-0238401P.
PR 06-OCT-2000; 2000US-0238402P.
PR 14-MAR-2001; 2001US-0275892P.
PR 08-JUN-2001; 2001US-0296860P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;
PI Rastelli L, Malyankar UM, Grosse WM, Alsobrook JP, Lepley DM;
PI Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;
PI Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;
XX
XX WPI; 2002-444103/47.
DR
DR P-PSDB; ABG61784.
XX
XX Novel isolated polypeptide, designated NOVX, useful for treating or
PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT metabolic, neurodegenerative, immune and hematopoietic disorders.
PT
XX
XX Claim 9; Page 62; 316pp; English.
XX
XX The invention describes an isolated polypeptide (I), useful in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease. (II), the polynucleotide encoding it (II), and an antibody
CC (III) to (I) are useful for treating or preventing cancer, metabolic
CC disorders, skin disorders, infectious disease, anorexia, behavioral
CC disorders, valve diseases, endocrine disorders, heart and blood
CC disorders, anxiety disorders, brain disorders, inflammatory disorders,
CC neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,
CC immune disorders, haematopoietic disorders, and the various
CC dyslipidaemias, metabolic disturbances associated with obesity, and the
CC metabolic syndrome X and wasting disorders associated with chronic
CC diseases. (I), (II) or (III) are useful in screening assays, detection
CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC predictive medicine (e.g., diagnostic assays and prognostic assays), and
CC in methods of treatment (e.g., therapeutic and prophylactic). (I) is
CC useful as immunogen to produce antibodies immunospecific for (I), to
CC screen for potential agonist and antagonist compounds, and as bait
CC protein in a two-hybrid or three-hybrid assay. (II) is useful in gene
CC therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)
CC gene, and to modulate activity of (I). A cell containing a vector
CC expressing (I) is useful for producing non-human transgenic animals. This
CC sequence encodes a novel human polypeptide described in the invention
XX
XX Sequence 1734 BP; 358 A; 516 C; 578 G; 282 T; 0 U; 0 Other;
SQ
Query Match 99.9%; Score 1576.4; DB 6; Length 1734;
Best Local Similarity 99.9%; Pred. No. 3.3e-311;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCCACCTTGTCTCGCAGCAAGCTGTCCACGTGGCCACGTCCTGTCCACCAAGTCC 60
DB 1623 ATGGCCACCTTGTCTCGCAGCAAGCTGTCCACGTGGCCACGTCCTGTCCACCAAGTCC 1564
QY 61 CAGGCCAAGATGAGCGGCATGTTCCGCCAGGATGGGTTTTCAGCGCGCCACGGATGAGGAG 120
DB 1563 CAGGCCAAGATGAGCGGCATGTTCCGCCAGGATGGGTTTTCAGCGCGCCACGGATGAGGAG 1504
QY 121 GCGGTGGGCTTCGCGCATTTGCCAGCACCTTCGACTTTTTCAGCGCGCCACGGATGAGGAG 180
DB 1503 GCGGTGGGCTTCGCGCATTTGCCAGCACCTTCGACTTTTTCAGCGCGCCACGGATGAGGAG 1444
QY 181 GACATCTGAAAGCCGAGGAGAGCCCTCGCGGGACGAGGCGCTGTAAGCGCCGCTCGAG 240
DB 1443 GACATCTGAAAGCCGAGGAGAGCCCTCGCGGGACGAGGCGCTGTAAGCGCCGCTCGAG 1384
QY 241 GGAGACATCATTTATCAGCGAGGCGAGCGAGTCTCTTCCCGCCCTCCGCTCCAAAGGAC 300

Db 1394 GGCAGCCTCAGCGGCGCGGCTCTGTTCTTGTGCTCCAGCCTCTTTCACCTGCGCCTG 1453
Qy 1441 CTCTGGCGCAAGCTGCTGGGACCAAGTCTTCTTCGAGTGGCATCTTCTGTCATCGGC 1500
Db 1454 CTCTGGCGCAAGCTGCTGGGACCAAGTCTTCTTCGAGTGGCATCTTCTGTCATCGGC 1513
Qy 1501 GGCATCTGCAAGCTGCTGGGCTTCTGTCACCTCCCTCGAGGGCTCATCGAAGCCTACCGA 1560
Db 1514 GGCATCTGCAAGCTGCTGGGCTTCTGTCACCTCCCTCGAGGGCTCATCGAAGCCTACCGA 1573
Qy 1561 ACCAAGCGGAGGACTAG 1578
Db 1574 ACCAAGCGGAGGACTAG 1591

RESULT 3

ABK12349

ID ABK12349 standard; DNA; 2261 BP.

XX ABK12349;

DT 18-JUN-2002 (first entry)

XX Human gamma-aminobutyric acid (GABA) transporter DNA sequence.

XX Human; gamma-aminobutyric acid transporter; GABA transporter; ds;

KW Gene therapy; Alzheimer's disease; Parkinson's disease; epilepsy.

XX Homo sapiens.

XX WO200220753-A2.

XX 14-MAR-2002.

XX 28-AUG-2001; 2001WO-US026778.

XX 01-SEP-2000; 2000US-0230178P.

XX (LEXI-) LEXICON GENETICS INC.

XX Friddle CJ, Gerhardt B, Hu Y;

XX WPI; 2002-292265/33.

XX New human nucleic acid, useful e.g. in drug screening, encodes gamma-

PT aminobutyric acid transporter protein.

XX Disclosure; Page 36; 36pp; English.

XX The present invention relates to a new nucleic acid that encodes a fully
CC defined 525 amino acid sequence, as given in the specification, or
CC hybridizes under stringent conditions to a fully defined 1578 base pair
CC sequence, or its complement, as given in the specification. The invention
CC is useful for genomic mapping or identifying and validating exon splice
CC junctions, in molecular mutagenesis and evolution of proteins, for
CC recombinant production of encoded proteins, for production of transgenic
CC knock-out or knock-in animals useful for identifying phenotypic
CC expression of the nucleic acid of the invention and for assigning
CC function to unknown genes, and also for generating antibodies to
CC homologous or orthologous proteins. Other uses of the invention include
CC as primers/probes for cloning, sequencing and library screening (for drug
CC discovery), to characterise temporal and tissue-specific expression of
CC genes, for identifying mutations and as a source of antisense, double-
CC stranded or ribozyme reagents for gene therapy. The molecules of the
CC invention and their fragments are used to raise antibodies, to identify
CC related gene products, to screen for potential therapeutic agents, e.g.
CC for treatment of Alzheimer's or Parkinson's diseases or epilepsy, i.e.
CC (antagonists or other modulators), and as therapeutic agents. The present
CC nucleic acid sequence represents the human gamma-aminobutyric acid (GABA)
CC transporter DNA sequence of the invention. Human GABA transporters share
CC structural similarity with mammalian GABA transporters

SQ Sequence 2261 BP; 348 A; 781 C; 680 G; 452 T; 0 U; 0 Other;
Query Match 100.0%; Score 1578; DB 6; Length 2261;
Best Local Similarity 100.0%; Pred. No. 1.6e-311;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCCACCTTGTCTCCGACGCAAGCTGTCCAAACGTGGCCACGTCCTGTCCTCAACAAAGTCC 60
Db 409 ATGGCCACCTTGTCTCCGACGCAAGCTGTCCAAACGTGGCCACGTCCTGTCCTCAACAAAGTCC 468
Qy 61 CAGGCCAAGATGAGCGGCATGTTGCCAGAGATGGGTTTTCAGGCGGCCACGGATAGGAG 120
Db 469 CAGGCCAAGATGAGCGGCATGTTGCCAGAGATGGGTTTTCAGGCGGCCACGGATAGGAG 528
Qy 121 CGGTGGGCTTTCGGCATTTGCGACACCTTTCAGACACCTTTCAGACACCGCCAGGGCTTCAGATG 180
Db 529 CGGTGGGCTTTCGGCATTTGCGACACCTTTCAGACACCTTTCAGACACCGCCAGGGCTTCAGATG 588
Qy 181 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGACAGAGGGCGCTGAAAGCGCCCTCGAG 240
Db 589 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGACAGAGGGCGCTGAAAGCGCCCTCGAG 648
Qy 241 GGAGACATCCATTTATCAGCGAGGAGCGGAGTCTCTTCCGCGCTTCAGGCTTCAAGGAC 300
Db 649 GGAGACATCCATTTATCAGCGAGGAGCGGAGTCTCTTCCGCGCTTCAGGCTTCAAGGAC 708
Qy 301 CAGGTGGGAGTGGTGGGCAATTTCGGGGCCACAGCAAGCCGCTGAAATCAGCGCGTGGGAG 360
Db 709 CAGGTGGGAGTGGTGGGCAATTTCGGGGCCACAGCAAGCCGCTGAAATCAGCGCGTGGGAG 768
Qy 361 GCAGGCTGGAACGCTGACCAAGCCATCCAGGGCATGTTTCGTCTGGGCTTACCTTACGCC 420
Db 769 GCAGGCTGGAACGCTGACCAAGCCATCCAGGGCATGTTTCGTCTGGGCTTACCTTACGCC 828
Qy 421 ATCTGACAGCGCGCTACCTGGGGTTGTTTCTCATCTTCGCGCGCTTGTGTGCTGC 480
Db 829 ATCTGACAGCGCGCTACCTGGGGTTGTTTCTCATCTTCGCGCGCTTGTGTGCTGC 888
Qy 481 TACACCGGCAGATCTCATCGCGTGCCTGACGAGGAGATGAAAGCGGCGAGTGGTG 540
Db 889 TACACCGGCAGATCTCATCGCGTGCCTGACGAGGAGATGAAAGCGGCGAGTGGTG 948
Qy 541 CGCGTGGCGGACTCGTACGTGGCCATAGCCAAACGCTCTGTCGCGCCCGCGCTTCCCAAGC 600
Db 949 CGCGTGGCGGACTCGTACGTGGCCATAGCCAAACGCTCTGTCGCGCCCGCGCTTCCCAAGC 1008
Qy 601 CTGGCGGCCCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGTGATGACGTGCATCTCTG 660
Db 1009 CTGGCGGCCCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGTGATGACGTGCATCTCTG 1068
Qy 661 TACGTGGTGTGAGTGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCGTGTCCGAG 720
Db 1069 TACGTGGTGTGAGTGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCGTGTCCGAG 1128
Qy 721 AAGTCTCTGGTCCATTATCGCCACGCGCGCTGCTGCTCTGCGCTTCTTAAAGAACCTC 780
Db 1129 AAGTCTCTGGTCCATTATCGCCACGCGCGCTGCTGCTCTGCGCTTCTTAAAGAACCTC 1188
Qy 781 AAGCCGCTGTCCAAAGTTCAGTCTGCTGTGCACTCTGGGCCACTTCGTCAATCAATATCTG 840
Db 1189 AAGCCGCTGTCCAAAGTTCAGTCTGCTGTGCACTCTGGGCCACTTCGTCAATCAATATCTG 1248
Qy 841 GTCATAGCTACTGTCTATCGCGGGCGCGGACTGGGCTGGGAGAAAGTCAAGTCTTAC 900
Db 1249 GTCATAGCTACTGTCTATCGCGGGCGCGGACTGGGCTGGGAGAAAGTCAAGTCTTAC 1308
Qy 901 ATCGACGTCAAGAAGTTCCTCCATTCGATTCGATTCATCTGCTTTCAGCTTACACGCTCTAG 960
Db 1309 ATCGACGTCAAGAAGTTCCTCCATTCGATTCGATTCATCTGCTTTCAGCTTACACGCTCTAG 1368
Qy 961 ATCTTCTGCTCTGCTGGAGGCAATATGACAGACCCAGCGAGTTCACCTGCATGATG 1020
Db 1369 ATCTTCTGCTCTGCTGGAGGCAATATGACAGACCCAGCGAGTTCACCTGCATGATG 1428

KW candidate nucleus; amygdala; thalamus; psychiatric disorder; anxiety;
KW neurological disorder; neurodegenerative disorder; bipolar disorder;
KW unipolar disorder; schizophrenia; psychotic disorder; ss.
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 14..1591
FT /*tag= a
FT /product= "neurotransmitter polypeptide HIPHUM 0000057"
XX
XX GB2366566-A.
XX
XX 13-MAR-2002.
XX
XX 31-MAY-2001; 2001GB-00013257.
XX
XX 01-JUN-2000; 2000GB-00013239.
XX
XX (GLAX) GLAXO GROUP LTD.
XX
XX Sala CF, Terstappen GC;
XX WPI; 2002-397154/43.
XX P-PSDB; ABB09876.
XX
XX Isolated neurotransmitter polypeptide HIPHUM 0000057, useful as a drug
XX target for treating neurological/neurodegenerative disorders, e.g.
XX bipolar disorder, unipolar disorder, anxiety, schizophrenia and/or
XX psychotic disorders.
XX Claim 5; Page 23-26; 35pp; English.
XX
XX The present sequence encodes a human neurotransmitter polypeptide,
CC designated HIPHUM 0000057. HIPHUM 0000057 is mapped to chromosomal region
CC 20q12-q13. HIPHUM 0000057 is specifically expressed in the brain, and the
CC highest level of expression is in the candidate nucleus. High levels of
CC expression are also found in the amygdala and thalamus. These areas of the
CC brain are associated with psychiatric, neurological and neurodegenerative
CC disorders. The neurotransmitter transporter polypeptide HIPHUM 0000057 is
CC a screening target for identification and development of drugs, i.e.
CC modulators of neurotransmitter transporter polypeptide expression and/or
CC activity. These drugs may be used in the prevention and/or treatment of
CC psychiatric disorders or neurological/neurodegenerative disorders, e.g.
CC bipolar disorder, unipolar disorder, anxiety, schizophrenia and/or
CC psychotic disorders
XX
XX Sequence 1595 BP; 268 A; 521 C; 480 G; 326 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1578; DB 6; Length 1595;
Best Local Similarity 100.0%; Pred. No. 1.5e-311;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTGCTCCGACGAAAGTGTCCAAAGTGGCCACGTCCTGTCACCAAGTCC 60
DB 14 ATGGCCACCTTGCTCCGACGAAAGTGTCCAAAGTGGCCACGTCCTGTCACCAAGTCC 73

QY 61 CAGGCCAAGATAGCGGCATGTTGCGCAGATGGGTTTCAGCGCGCCACGAGTAGGAG 120
DB 74 CAGGCCAAGATAGCGGCATGTTGCGCAGATGGGTTTCAGCGCGCCACGAGTAGGAG 133

QY 121 GCGGTGGGCTTCGCGCATTCGCGACGACCTCGACTTTGAGCACCGCCAGGCGCTGCAGATG 180
DB 134 GCGGTGGGCTTCGCGCATTCGCGACGACCTCGACTTTGAGCACCGCCAGGCGCTGCAGATG 193

QY 181 GACATCCTGAAAGCCGAGGAGAGCCCTGCGCGGAGCGAGGCGCTGAAGCGCCGCTCGAG 240
DB 194 GACATCCTGAAAGCCGAGGAGAGCCCTGCGCGGAGCGAGGCGCTGAAGCGCCGCTCGAG 253

QY 241 GGAGACATTCATTATCAGCGAGGAGCGAGCTCTCTTGCCTCCGCTCCAGGAC 300
DB 254 GGAGACATTCATTATCAGCGAGGAGCGAGCTCTCTTGCCTCCGCTCCAGGAC 313

QY 301 CAGGTGGAGTGTGTGGCGAATTCGGGGGCCACGACAAAGCCAAATACCGCGTGGAG 360
DB 314 CAGGTGGAGTGTGTGGCGAATTCGGGGGCCACGACAAAGCCAAATACCGCGTGGAG 373

QY 361 GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGTGTGGGCTTACCCCTAGGCC 420
DB 374 GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGTGTGGGCTTACCCCTAGGCC 433

QY 421 ATCTCTGACCGCGGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGCTTGTGTGTGC 480
DB 434 ATCTCTGACCGCGGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGCTTGTGTGTGC 493

QY 481 TACACCGCAAGATCTCATCGGTGCTGTACGAGGAGAAATGAAGACGCGAGGTGCTG 540
DB 494 TACACCGCAAGATCTCATCGGTGCTGTACGAGGAGAAATGAAGACGCGAGGTGCTG 553

QY 541 CGCGTGGGACCTCGTACGTGGCCATAGCAACGCCCTGTGCGGCCCGCGCTTCCCAACG 600
DB 554 CGCGTGGGACCTCGTACGTGGCCATAGCAACGCCCTGTGCGGCCCGCGCTTCCCAACG 613

QY 601 CTGGGCGCGGAGTGTGTAACGTAGCGCAGATCATCGAGCTGTGTGATGACGTGATCCTG 660
DB 614 CTGGGCGCGGAGTGTGTAACGTAGCGCAGATCATCGAGCTGTGTGATGACGTGATCCTG 673

QY 661 TAGCTGTGTGAGTGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCGCTGTGCGAG 720
DB 674 TAGCTGTGTGAGTGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCGCTGTGCGAG 733

QY 721 AAGTCTCTGTGTCATTTATCGCACGGCGCTGCTGCTGCTTGCCTTAAAGAACCTC 780
DB 734 AAGTCTCTGTGTCATTTATCGCACGGCGCTGCTGCTGCTTGCCTTAAAGAACCTC 793

QY 781 AAGCGCGTGTCAAGTTCAGTCTGTGTGCACTCTTGGGCCACTTCTGTCATCATATCTG 840
DB 794 AAGCGCGTGTCAAGTTCAGTCTGTGTGCACTCTTGGGCCACTTCTGTCATCATATCTG 853

QY 841 GTCATAGCTACTGTCTATCGCGGGCGCGACTGGGCTGGGAGAGGTCAAGTTCATC 900
DB 854 GTCATAGCTACTGTCTATCGCGGGCGCGACTGGGCTGGGAGAGGTCAAGTTCATC 913

QY 901 ATCGAGCTCAAGAAAGTTCCTCCATCTCATTTGGCATCATGTTTTCAGTACACGTCTCAG 960
DB 914 ATCGAGCTCAAGAAAGTTCCTCCATCTCATTTGGCATCATGTTTTCAGTACACGTCTCAG 973

QY 961 ATCTTCTGCTGCTTGTGAGGCAATATGACAGCCAGGAGTTCATGATGATG 1020
DB 974 ATCTTCTGCTGCTTGTGAGGCAATATGACAGCCAGGAGTTCATGATGATG 1033

QY 1021 AACTGACGCACATCCGACGCTGCGTGTCTCAAGGGGCTCTTCCGGCTCGTCCGCTACCTC 1080
DB 1034 AACTGACGCACATCCGACGCTGCGTGTCTCAAGGGGCTCTTCCGGCTCGTCCGCTACCTC 1093

QY 1081 ACCTGGCGCGACGAGACCAAGAGGTTCATACCGATAAAGTTCCTCCGCGCTCCATCCGGGCC 1140
DB 1094 ACCTGGCGCGACGAGACCAAGAGGTTCATACCGATAAAGTTCCTCCGCGCTCCATCCGGGCC 1153

QY 1141 GTGGTCAACATCTTTCTGTGTGCGCAAGGCGCTGTTGTCTATTCCTCTGCGCATTTCTTGGC 1200
DB 1154 GTGGTCAACATCTTTCTGTGTGCGCAAGGCGCTGTTGTCTATTCCTCTGCGCATTTCTTGGC 1213

QY 1201 GCTGTGAGGTGCTGAGAGTCTGCTTCCAGGAGGAGGAGCGCGGCTTTTCCCGGCC 1260
DB 1214 GCTGTGAGGTGCTGAGAGTCTGCTTCCAGGAGGAGGAGCGCGGCTTTTCCCGGCC 1273

QY 1261 TGCTACAGCGCGGACCGGCGGCTGAAAGTCTTGGGGGCTGACGCTGCGGCTGCGGCTCGTC 1320
DB 1274 TGCTACAGCGCGGACCGGCGGCTGAAAGTCTTGGGGGCTGACGCTGCGGCTGCGGCTCGTC 1333

QY 1321 GTCTTACGCTGCTCATGCGCAATTTATGTGCGCACTTCGCGCTGTCTATGCGGCTCAACC 1380
DB 1334 GTCTTACGCTGCTCATGCGCAATTTATGTGCGCACTTCGCGCTGTCTATGCGGCTCAACC 1393

QY 1381 GCGAGCTCACGGGCGCGGCTCTGTGTTCTTGTGCTGCCAGGCTCTTTTACCTGCGGCTG 1440

sequence, or its complement, as given in the specification. The invention is useful for genomic mapping or identifying and validating exon splice junctions, in molecular mutagenesis and evolution of proteins, for recombinant production of encoded proteins, for production of transgenic knock-out or knock-in animals useful for identifying phenotypic expression of the nucleic acid of the invention and for assigning function to unknown genes, and also for generating antibodies to homologous or orthologous proteins. Other uses of the invention include as primers/probes for cloning, sequencing and library screening (for drug discovery), to characterise temporal and tissue-specific expression of genes, for identifying mutations and as a source of antisense, double-stranded or ribozyme reagents for gene therapy. The molecules of the invention and their fragments are used to raise antibodies, to identify related gene products, to screen for potential therapeutic agents, e.g. for treatment of Alzheimer's or Parkinson's diseases or epilepsy, i.e. (ant)agonists or other modulators, and as therapeutic agents. The present nucleic acid sequence encodes the human gamma-aminobutyric acid (GABA) transporter protein of the invention. The protein shares structural similarity with mammalian GABA transporters.

Sequence 1578 BP; 267 A; 510 C; 475 G; 326 T; 0 U; 0 Other;

Query Match 100.0%; Score 1578; DB 6; Length 1578;

Best Local Similarity

Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCCACCTTGCCTCGCAGCAAGCTGTCTCAACGTGGCCACCTCGTGTCTCAACAAGATCC	600
Db	1	ATGGCCACCTTGCCTCGCAGCAAGCTGTCTCAACGTGGCCACCTCGTGTCTCAACAAGATCC	60
Qy	61	CAGGCCAAGATCAGCGGGATGTTCCCAAGGATGGTGTTCAGGCGGCCAGGATGAGGAG	120
Db	61	CAGGCCAAGATCAGCGGGATGTTCCCAAGGATGGTGTTCAGGCGGCCAGGATGAGGAG	120
Qy	121	GGGTGGGCTTCGGGCATTTGGCA CGACTCTGACTTTTGGACACCGCCAGGCGCTTCAGATG	180
Db	121	GGGTGGGCTTCGGGCATTTGGCAGACCTGACTTTTGGACACCGCCAGGCGCTTCGAGATG	180
Qy	181	GACATCTCAAAACCGAGGGAGAGCCCTGCGGGACAGAGGCGCTGTAAGGCCCGCTCGAG	240
Db	181	GACATCTCAAAACCGAGGGAGAGCCCTGCGGGACAGAGGCGCTGTAAGGCCCGCTCGAG	240
Qy	241	GGAGACATCCATTATCAGCGAGCAGCGGAGCTCTCTTGGCGCCCTCCGGCTCCAAAGAC	300
Db	241	GGAGACATCCATTATCAGCGAGCAGCGGAGCTCTCTTGGCGCCCTCCGGCTCCAAAGAC	300
Qy	301	CAGGTGGGAGGTGGTGGCGAATTTCCGGGGCCACAGCAAGCCCAAAATCAGCGGTGGAG	360
Db	301	CAGGTGGGAGGTGGTGGCGAATTTCCGGGGCCACAGCAAGCCCAAAATCAGCGGTGGAG	360
Qy	361	GCAGGCTGGAACGTGACCAAGCCCATCAGGGGCATGTTCTGTGTGGGCTTACCTTACGCC	420
Db	361	GCAGGCTGGAACGTGACCAAGCCCATCAGGGGCATGTTCTGTGTGGGCTTACCTTACGCC	420
Qy	421	ATCCTGCACGGCGGTACCTGGGGTGTGTTCTCATCATCTTCGCGCGGTTGTGTGCTGC	480
Db	421	ATCCTGCACGGCGGTACCTGGGGTGTGTTCTCATCATCTTCGCGCGGTTGTGTGCTGC	480
Qy	481	TACACCGGCAAGATCCTCATCCGTGCTGTACAGGAGAAATGAAGACGCGGAGGTGGTG	540
Db	481	TACACCGGCAAGATCCTCATCCGTGCTGTACAGGAGAAATGAAGACGCGGAGGTGGTG	540
Qy	541	CGGTGGGGAATCGTATCGTGCCATAGCCAA CGCTCTGCGGCCCGCGGCTTCCCAACG	600
Db	541	CGGTGGGGAATCGTATCGTGCCATAGCCAA CGCTCTGCGGCCCGCGGCTTCCCAACG	600
Qy	601	CTGGGCGGCGGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGATGAGTGCATCTCTG	660
Db	601	CTGGGCGGCGGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGATGAGTGCATCTCTG	660
Qy	661	TACGTGGTGGTGGCAACCTCATGTACAAAGCTTCCGGGGCTGCCCGTGTCCGAG	720
Db	661	TACGTGGTGGTGGCAACCTCATGTACAAAGCTTCCGGGGCTGCCCGTGTCCGAG	720

RESULT 2

ABL56565

ID ABL56565 standard; cDNA; 1595 BP.

XX

AC ABL56565;

XX

DT 05-AUG-2002 (first entry)

XX

DE Nucleotide sequence of neurotransmitter polypeptide HIPHUM 0000057.

XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 18:26:25 ; Search time 750 Seconds
(without alignments)
11044.782 Million cell updates/sec

Title: US-09-940-919-1
Perfect score: 1578
Sequence: 1 atggccacctgtctccgag.....gaaccaacggaggactag 1578

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	6	ABK12348
2	1578	100.0	1595	6	ABL56565
3	1578	100.0	2261	6	ABK12349
4	1576.4	99.9	1734	6	ABK32051
5	1576.4	99.9	1763	6	ABK32049
6	1576.4	99.9	1800	5	AAS15704
7	1574.8	99.8	1734	6	ABK32050
8	1574.8	99.8	1734	6	ABK32052
9	1574.8	99.8	1734	6	ABK32053
10	1574.8	99.8	1734	6	ABK32054
11	1344.2	85.2	1648	5	AAS06061
12	1140.2	72.2	1723	5	AAS06087
13	387.2	24.5	1095	5	AAS81399
14	366	23.2	1650	4	ABL14045
15	366	23.2	3722	4	ABL14044
16	237	15.0	2255	9	ACH15349
17	232.2	14.7	2740	12	ADQ08701
18	177.2	11.2	498	9	ACH38839
19	131	8.3	448	5	ABA20103
20	101.8	6.5	451	5	ABA13679
21	77.2	4.9	1580	8	ADA70336

ALIGNMENTS

RESULT 1

ABK12348

ID ABK12348 standard; cDNA; 1578 BP.

XX

AC ABK12348;

DT 18-JUN-2002 (first entry)

XX

DE cDNA encoding human gamma-aminobutyric acid (GABA) transporter protein.

XX Human; gamma-aminobutyric acid transporter; GABA transporter; gene; ss;

KW gene therapy; Alzheimer's disease; Parkinson's disease; epilepsy.

XX Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..1578

FT /*tag= a

FT /product= "Human gamma-aminobutyric acid (GABA)

FT transporter protein"

XX

XX WO200220753-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX


```
QY 421 ATCTGCGAGCGGCTACTCGGGTTGTTCTCATCATCTTCGCCGCGTGTGTGCTGC 480
Db 421 ATCTGCGAGCGGCGCTACTCGGGTTGTTCTCATCATCTTCGCCGCGTGTGTGCTGC 480
QY 481 TACACCGGCAAGATCCTCATCGCGTGCCTGTACGAGGAGAAATGAAGCGGAGGTGGTG 540
Db 481 TACACCGGCAAGATCCTCATCGCGTGCCTGTACGAGGAGAAATGAAGCGGAGGTGGTG 540
QY 541 CGCGTGGGACTCTGTAGTGCCCATAGCCAAAGCGCTGTGCGCCCGCGCTTCCCAACG 600
Db 541 CGCGTGGGACTCTGTAGTGCCCATAGCCAAAGCGCTGTGCGCCCGCGCTTCCCAACG 600
QY 601 CTGGGCGGCCGAGTGTGAACCTAGCGAGATCATCGAGCTGTGATGACGTGCATCTCTG 660
Db 601 CTGGGCGGCCGAGTGTGAACCTAGCGAGATCATCGAGCTGTGATGACGTGCATCTCTG 660
QY 661 TACGTGGTGGTGAAGTGGCAACCTCATGTACAACAGCTTCCCGGGCTGCCCCGTGTCGAG 720
Db 661 TACGTGGT----- 668
QY 721 AAGTCTGTGTCATTATCGCCACGCGCGTGTGCTGCGCTTCCCTTAAGAACCTC 780
Db 669 -----GAACCTC 675
QY 781 AAGGCCGTGTCCAAGTTCAGTCTGTGTGCTGTGCCCCACTTCGTGATCAATATCTCTG 840
Db 676 AAGGCCGTGTCCAAGTTCAGTCTGTGTGCTGTGCCCCACTTCGTGATCAATATCTCTG 735
QY 841 GTATAGGCTACTGTCTATCGCGGCGCGCGACTGGGCTGGGAGAGGTCAAGTTCTAC 900
Db 736 GTATAGGCTACTGTCTATCGCGGCGCGCGACTGGGCTGGGAGAGGTCAAGTTCTAC 795
QY 901 ATCGAGCTCAAGAAGTTCCTCATTTGGCATCATCGTGTTCAGCTACACGTCTCAG 960
Db 796 ATCGAGCTCAAGAAGTTCCTCATTTGGCATCATCGTGTTCAGCTACACGTCTCAG 855
QY 961 ATCTTCTGCTTCTGCGGAGGCGCAATATGACAGCGCCAGCGAGTTCACCTGCAATGATG 1020
Db 856 ATCTTCTGCTTCTGCGGAGGCGCAATATGACAGCGCCAGCGAGTTCACCTGCAATGATG 915
QY 1021 AACTGGACGCACATCGAGCCTCGTGTCTCAAGGGCTCTTCGCGCTCGTGCCTACCTC 1080
Db 916 AACTGGACGCACATCGAGCCTCGTGTCTCAAGGGCTCTTCGCGCTCGTGCCTACCTC 975
QY 1081 ACCTGGGCGGACGAGACCAAGAGGTATACGAGTAACCTGCCGCGCTCATCGCGCC 1140
Db 976 ACCTGGGCGGACGAGACCAAGAGGTATACGAGTAACCTGCCGCGCTCATCGCGCC 1035
QY 1141 GTGGTCAACATCTTTCTGTGTGCCAAGCGCTGTGTCTATCTCTGCGCTTCTTTGCC 1200
Db 1036 GTGGTCAACATCTTTCTGTGTGCCAAGCGCTGTGTCTATCTCTGCGCTTCTTTGCC 1095
QY 1201 GCTGTGAGGTGTGGAGAAGTGTCTTTCCAGGAAGGAGCGCGCCCTTTTCCCGGCC 1260
Db 1096 GCTGTGAGGTGTGGAGAAGTGTCTTTCCAGGAAGGAGCGCGCCCTTTTCCCGGCC 1155
QY 1261 TGTACAGGGCGGACGCGGCGCTGAAGTCTGCGGGCTGACGCTGGCTGCGCGCTCGTC 1320
Db 1156 TGTACAGGGCGGACGCGGCGCTGAAGTCTGCGGGCTGACGCTGGCTGCGCGCTCGTC 1215
QY 1321 GTCTTCACGCTGTCTATGCGCATTTATGTGCGCACTTCGCGCTGTCTCATGGGCTCAC 1380
Db 1216 GTCTTCACGCTGTCTATGCGCATTTATGTGCGCACTTCGCGCTGTCTCATGGGCTCAC 1275
QY 1381 GGCAGCCTCACGGGCGCGGCTCTGTGTTCTTGTGCCAGGCTCTTTTCACTGCGCCTG 1440
Db 1276 GGCAGCCTCACGGGCGCGGCTCTGTGTTCTTGTGCCAGGCTCTTTTCACTGCGCCTG 1335
QY 1441 CTCTGGCGCAAGTGTGTGGACCAAGTCTTCTTCGAGCTGCGCATCTTGTCTATCGGC 1500
Db 1336 CTCTGGCGCAAGTGTGTGGACCAAGTCTTCTTCTCGAGCTGCGCATCTTGTCTATCGGC 1395
```

```
QY 1501 GGCACTCTGACGCTGTCCGGCTTCGTGCACTCCCTCCGAGGGCTTCATCGAAGCCTTACCGA 1560
Db 1396 GGCACTCTGACGCTGTCCGGCTTCGTGCACTCCCTCCGAGGGCTTCATCGAAGCCTTACCGA 1455
QY 1561 ACCAACGCGGAGGACTAG 1578
Db 1456 ACCAACGCGGAGGACTAG 1473
```

Search completed: November 3, 2004, 21:05:54
Job time : 6616 secs

QY	271	GCTCTCTGCGCCCTTCGGCTTCAAAGACAGGTGGAGAGTGTGGCGAAATTCGGGGGC	330
DB	121		
		GCTCTCTGCGCCCTTCGGCTTCAAAGACAGGTGGAGAGTGTGGCGAAATTCGGGGGC	180
QY	331	CACGACAAAGCCAAATACACGGCGTGGAGGACGCTGGAACGTGACCAAGCCATCCAG	390
DB	181		
		CACGACAAAGCCAAATACACGGCGTGGAGGACGCTGGAACGTGACCAAGCCATCCAG	240
QY	391	GGCATGTTCTGCTGGGCTTACCCTACGCCATCTCTGACGGCGGCTACTCTGGGGTTGTTT	450
DB	241		
		GGCATGTTCTGCTGGGCTTACCCTACGCCATCTCTGACGGCGGCTACTCTGGGGTTGTTT	300
QY	451	CTCATCATCTTTCGCGCGTGTGTGTCTACACCGCAAGATCTCTCATCGCTGCGCTG	510
DB	301		
		CTCATCATCTTTCGCGCGTGTGTGTCTACACCGCAAGATCTCTCATCGCTGCGCTG	360
QY	511	TACGAGGAGATGAAGACGGCGAGGTGTGCGGTGCGGACCTGTAAGTGGCCATAGCC	570
DB	361		
		TACGAGGAGATGAAGACGGCGAGGTGTGCGGTGCGGACCTGTAAGTGGCCATAGCC	420
QY	571	AACGCTCTGTCGCGCCCGCTTCCAAACGCTGGGCGCGAGTGGTGAAGCTAGCGCGAG	630
DB	421		
		AACGCTCTGTCGCGCCCGCTTCCAAACGCTGGGCGCGAGTGGTGAAGCTAGCGCGAG	480
QY	631	ATCATCGAGCTGGTGAATGACGTGATCTGTACGTGGTGGTGAAGTGGCAACCTCATGTAC	690
DB	481		
		ATCATCGAGCTGGTGAATGACGTGATCTGTACGTGGTGGTGAAGTGGCAACCTCATGTAC	540
QY	691	AACAGCTTCCGCGGCTGCGCGTGTGGAGAAAGTCTGGTCCATTAATGCGCAAGCGCGTG	750
DB	541		
		AACAGCTTCCGCGGCTGCGCGTGTGGAGAAAGTCTGGTCCATTAATGCGCAAGCGCGTG	600
QY	751	CTGCTGCCCTTGGCGCTTCTTAAAGAACTCAAGCGCGTGTCCAAAGTTCAGTCTGCTGTC	810
DB	601		
		CTGCTGCCCTTGGCGCTTCTTAAAGAACTCAAGCGCGTGTCCAAAGTTCAGTCTGCTGTC	660
QY	811	ACTCTGGCCACTTCGTTCATCAATATCTGTGTCATAGCTACTGTCTATCGCGGCGCGC	870
DB	661		
		ACTCTGGCCACTTCGTTCATCAATATCTGTGTCATAGCTACTGTCTATCGCGGCGCGC	720
QY	871	GACTGGGCTGGGAGAGGTCAAGTTCATACGACGTCGAAGAGTTCCTCCATCTCCATT	930
DB	721		
		GACTGGGCTGGGAGAGGTCAAGTTCATACGACGTCGAAGAGTTCCTCCATCTCCATT	780
QY	931	GGCATCATCTGTTTCAGCTACACGCTCAGATCTTCTGCTTCCGCTGGAGGGCAATATG	990
DB	781		
		GGCATCATCTGTTTCAGCTACACGCTCAGATCTTCTGCTTCCGCTGGAGGGCAATATG	840
QY	991	CAGCAGCCACGAGATTCCACTGCATGATGAATCGACGACATCGCAGCCTGGCTGCTC	1050
DB	841		
		CAGCAGCCACGAGATTCCACTGCATGATGAATCGACGACATCGCAGCCTGGCTGCTC	900
QY	1051	AAGGGCTTTCGCGCTGTCGCTACCTCAGTGGGCGGACGAGACCAAGGAGGTATC	1110
DB	901		
		AAGGGCTTTCGCGCTGTCGCTACCTCAGTGGGCGGACGAGACCAAGGAGGTATC	960
QY	1111	ACGATTAACCTGCGCGGCTCCATCGCGCGTGTCAACATCTTTCTGTGGGCCAAGCGG	1170
DB	961		
		ACGATTAACCTGCGCGGCTCCATCGCGCGTGTCAACATCTTTCTGTGGGCCAAGCGG	1020
QY	1171	CTGTTGTCTTCTCTGTGCCATTCTTTGCGCTGTGAGGTGCTGGAGAGTCTGCTCTC	1230
DB	1021		
		CTGTTGTCTTCTCTGTGCCATTCTTTGCGCTGTGAGGTGCTGGAGAGTCTGCTCTC	1080
QY	1231	CAGGAAGCAGCCGCGCTTTTCCGCGCTGTACAGCGCGACGGCGCTCAAGTCC	1290
DB	1081		
		CAGGAAGCAGCCGCGCTTTTCCGCGCTGTACAGCGCGACGGCGCTCAAGTCC	1140
QY	1291	TGGGGGCTGACGCTGGCGTGGCGCTGTGTCTTCAAGCTGCTCATGGCCATTTATGTG	1350
DB	1141		
		TGGGGGCTGACGCTGGCGTGGCGCTGTGTCTTCAAGCTGCTCATGGCCATTTATGTG	1200

Qy	1351	CCGCACCTTCGGCGCTGCTCATGAGGCGCTCACCGGAGAGCCTCAACGGGCGCGGCGCTCTGTGTTTC	1411	TTGCTGCGCCAGCCCTTTTCACTTCGGCGCTGCTCTGGCGCAAGCTGCTGTGGGACCAAGTC	1470
Db	1201	CCGCACCTTCGGCGCTGCTCATGAGGCGCTCACCGGAGAGCCTCAACGGGCGCGGCGCTCTGTGTTTC	1260	TTGCTGCGCCAGCCCTTTTCACTTCGGCGCTGCTCTGGCGCAAGCTGCTGTGGGACCAAGTC	1320
Qy	1411	TTGCTGCGCCAGCCCTTTTCACTTCGGCGCTGCTCTGGCGCAAGCTGCTGTGGGACCAAGTC	1470	TTCTTTGCAAGCTGCGCATCTTTCGTCATCGGCGGCATCTGCAGACGCTGTCGGGCTTCGTGCAC	1530
Db	1261	TTGCTGCGCCAGCCCTTTTCACTTCGGCGCTGCTCTGGCGCAAGCTGCTGTGGGACCAAGTC	1320	TTCTTTGCAAGCTGCGCATCTTTCGTCATCGGCGGCATCTGCAGACGCTGTCGGGCTTCGTGCAC	1380
Qy	1471	TTCTTTGCAAGCTGCGCATCTTTCGTCATCGGCGGCATCTGCAGACGCTGTCGGGCTTCGTGCAC	1530	TCCCTCGAGGGCCTCAFCGAAGCCTACCGAACCAACCGCGAGGACTAG	1578
Db	1321	TTCTTTGCAAGCTGCGCATCTTTCGTCATCGGCGGCATCTGCAGACGCTGTCGGGCTTCGTGCAC	1380	TCCCTCGAGGGCCTCATCGAAGCCTACCGAACCAACCGCGAGGACTAG	1428
Qy	1531	TCCCTCGAGGGCCTCAFCGAAGCCTACCGAACCAACCGCGAGGACTAG	1578		
Db	1381	TCCCTCGAGGGCCTCATCGAAGCCTACCGAACCAACCGCGAGGACTAG	1428		
RESULT 15					
Q0727095					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
ORIGIN					
Query Match					
Best Local Similarity					
Matches 1473; Conservative					
1	ATGCCACCTTGTCTCCGCGACGAAGCTGTCCAAAGTGGCCACGTCGTGTGCCAACAAAGTCC	60			
1	ATGCCACCTTGTCTCCGCGACGAAGCTGTCCAAAGTGGCCACGTCGTGTGCCAACAAAGTCC	60			
61	CAGGCCAAGATGAGCGGCATGTTTCGCCAGGATGGGTTTTCAGGCGGCCACGGATGAGGAG	120			
61	CAGGCCAAGATGAGCGGCATGTTTCGCCAGGATGGGTTTTCAGGCGGCCACGGATGAGGAG	120			
121	CGCGTGGGCTTCGCGCATTTGCGACGACCTCGACTTTTGAGCACCGCCAGGCGCTTCGAGATG	180			
121	CGCGTGGGCTTCGCGCATTTGCGACGACCTCGACTTTTGAGCACCGCCAGGCGCTTCGAGATG	180			
181	GACATCCTGAAAGCCGAGGGAGAGCCCTCGCGGGACGAGGGCGCTGAAAGCGCCGCTCGAG	240			
181	GACATCCTGAAAGCCGAGGGAGAGCCCTCGCGGGACGAGGGCGCTGAAAGCGCCGCTCGAG	240			
241	GGAGACATCATTTATCAGCGAGGCGGAGCTCTCTGCGGCCCTTCGGCTTCGAGGAC	300			
241	GGAGACATCATTTATCAGCGAGGCGGAGCTCTCTGCGGCCCTTCGGCTTCGAGGAC	300			
301	CAGGTGGGAGTGGTGGCGAATTCGCGGGGCCACGACAAAGCCCAAATATCACGGCGTGGGAG	360			
301	CAGGTGGGAGTGGTGGCGAATTCGCGGGGCCACGACAAAGCCCAAATATCACGGCGTGGGAG	360			
361	GCAGGCTGGAAACGTGACCAACCGCATCTCAGGCGCATGTTCTGCTGGGCTTACCTACGCC	420			
361	GCAGGCTGGAAACGTGACCAACCGCATCTCAGGCGCATGTTCTGCTGGGCTTACCTACGCC	420			

Db 1576 GTGGTCAACATCTTTCTGGTGCCCAAGGGCTGTGTCTCTATCTCTCTGCCATCTTTCGCC 1635
Qy 1201 GCTGTCGAGGTCCTGAGAGTCGCTCTTTCAGGAGGAGGAGCGCGCTTTTCCCGGCC 1260
Db 1636 GCTGTCGAGGTCCTGAGAGTCGCTCTTTCAGGAGGAGGAGCGCGCTTTTCCCGGCC 1695
Qy 1261 TGCTACAGCGGCGAGCGCGCTGAAGTCTTGGGGGCTGAGCGCTGCGCTGCGCGCTGCTC 1320
Db 1696 TGCTACAGCGGCGAGCGCGCTGAAGTCTTGGGGGCTGAGCGCTGCGCTGCGCGCTGCTC 1755
Qy 1321 GTCCTACGCTGCTCATGCGCATTTATGTGCGCGCATCTTTCGAGCGCTGCTCATGCGCGCTCACC 1380
Db 1756 GTCCTACGCTGCTCATGCGCATTTATGTGCGCGCATTTTCGAGCGCTGCTCATGCGCGCTCACC 1815
Qy 1381 GGCAGCCTCAGCGGCGCGCGCTGTTTCTGTCGCCCGCTCTTTCGAGCGCTGCTCATGCGCGCTG 1440
Db 1816 GGCAGCCTCAGCGGCGCGCGCTGTTTCTGTCGCCCGCTCTTTCGAGCGCTGCTCATGCGCGCTG 1875
Qy 1441 CTCTGGCGCAAGCTGCTGTGGCACCAGTCTTCTTCGAGCGCTGCGCATCTTTCGATCGGC 1500
Db 1876 CTCTGGCGCAAGCTGCTGTGGCACCAGTCTTCTTCGAGCGCTGCGCATCTTTCGATCGGC 1935
Qy 1501 GGCATCTGCGAGGTCGCGCTTCTGTCGACCTCCCTCGAGGCGCTCATGAGGCTTACCGA 1560
Db 1936 GGCATCTGCGAGGTCGCGCTTCTGTCGACCTCCCTCGAGGCGCTCATGAGGCTTACCGA 1995
Qy 1561 ACCACGCGGAGGACTAG 1578
Db 1996 ACCACGCGGAGGACTAG 2013

RESULT 14

BC036458 2174 bp mRNA linear PRI 30-SEP-2003
LOCUS Homo sapiens vesicular inhibitory amino acid transporter, mRNA
DEFINITION (CDNA clone IMAGE:4813726), partial cds.
ACCESSION BC036458
VERSION
KEYWORDS
SOURCE BC036458.2 GI:34785015
ORGANISM Homo sapiens (human)

REFERENCE

1 (bases 1 to 2174)
AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
PUBMED 12477932
2 (bases 1 to 2174)
AUTHORS Strausberg R.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK
COMMENT

USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Sep 16, 2003 this sequence version replaced gi:22209088.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadanesystemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 46 Row: d Column: 1.

FEATURES

source

1..2174
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4813726"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="vector: pBluescript"
<1..2174
/gene="VIAAT"
/note="synonym: ba12201.1"
/db_xref="LocusID:140679"
<1..1428
/gene="VIAAT"
/codon_start=1
/product="VIAAT protein"
/protein_id="AAH36458.2"
/db_xref="GI:34785016"
/db_xref="LocusID:140679"
/translations="GFSEHROGLMDILKAEGPCGDEAGAEAPVEGDIHVORGSGAPLP
PSGKQVGGGEGFGHDKPKITAEAGNVTNAIQNFVLGPAHLGGVGLGFLI
IFAAVCCYCTGKILACLYEENEDGVRVDSVAIANACAPFPFPLGGRGVNVAQ
IELVMTCLIVVYVGNLMYNSFPGLPVSKQSWIIATAVLLPCAFKLNLRVSKFSL
LCTLAHFVNIILVIAVCLSRARDWAEKVFYIDVKFPISIGIIVFYISQIFPLSL
EGNQKQSEPHCMNWTIIAACVLKGLFALVAILTWADETKEVITDNLPGSTRVAVNI
FLVAKALLSPFPFAAVVLEKSLFQBSRAFPACYSYDGRKLSWGLTLRCALVVF
TLVIAIYVHPFALMLGTLGAGLCLFLLSLFHLRLRLKLLWHQVFPFVAIFVIG
GICSVSGFVHSLEGLIEAYRTNAED"
193..1398
/note="VIAAT"
/gene="VIAAT"; Region: Transmembrane amino acid transporter protein. This transmembrane region is found in many amino acid transporters including UNC-47 and WTR. UNC-47 encodes a vesicular amino butyric acid (GABA) transporter, (VGAT). UNC-47 is predicted to have 10 transmembrane domains. MTR is a N system amino acid transporter system protein involved in methyltryptophan resistance. Other members of this family include proline transporters and amino acid permeases"
/db_xref="CDD:pfam01490"

gene

CDS

misc_feature

ORIGIN

Query Match 90.3%; Score 1424.8; DB 9; Length 2174;
Best Local Similarity 99.9%; Pred. No. 2.4e-206;
Matches 1426; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 151 GACTTTGAGCACCGCCAGGCGCTGCAGATGCACATCTCTGAAAGCCGAGGAGAGCCCTGC 210
Db 1 GGCCTTTGAGCACCGCCAGGCGCTGCAGATGCACATCTCTGAAAGCCGAGGAGAGCCCTGC 60
Qy 211 GGGGACGAGGCGCTGAAGCGCCCTCGAGGAGACATCCATTATCAGGAGGAGCGGGA 270
Db 61 GGGGACGAGGCGCTGAAGCGCCCTCGAGGAGACATCCATTATCAGGAGGAGCGGGA 120

KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Macaca fascicularis (crab-eating macaque)
ORGANISM	Macaca fascicularis
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
TITLE	Cercopitheciae; Macaca.
JOURNAL	1 (sites)
REFERENCE	Oada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M.,
AUTHORS	Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.
TITLE	Isolation of full-length cDNA clones from macaque brain cDNA
JOURNAL	libraries
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 2756)
TITLE	Hashimoto, K., Oada, N., Hida, M., Kusuda, J. and Sugano, S.
JOURNAL	Direct Submission
REFERENCE	Submitted (11-JUN-2001) Katsuyuki Hashimoto, National Institute of
AUTHORS	Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
TITLE	1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
JOURNAL	(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,
COMMENT	Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)
Lab host:	TOP10
Vector:	pME18S-FL3 (Acc.No. AB009864)
R. Site1:	DrailII (CACGTGTGG)
R. Site2:	DrailII (CACCATGTG)
Description:	1st strand cDNA was primed with an oligo(dT) primer [ATGGGCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DrailII sites of pME18S-FL3. XhoI sites just outside the DrailII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method (Sugano et al., , Institute of Medical Science, University of Tokyo).
Custom primer used for sequencing	
(5' end primer [CTTCTGCTCTAAAGCTGCG];	
3' end primer [CGACCTGCAGTCGAGCAC];	
FEATURES	Location/Qualifiers
source	1..2756
	/organism="Macaca fascicularis"
	/mol_type="mRNA"
	/db_xref="taxon:9541"
	/clone="OccE-21148"
	/sex="male"
	/tissue_type="cerebellum cortex"
	/clone_lib="macaque brain cDNA library QccB"
	/dev_stage="adult"
	436..2013
CDS	/codon_start=1
	/product="hypothetical protein"
	/protein_id="BAB60726.1"
	/db_xref="GI:114388326"
	/translation="MATLRSKLSNVATSVNSQAKVSGMFARMGFQATDEEAVGF AHCDJDFHROQLMDILKAEPCGDEGAPVEGDHLYORGSAPLPPSGSKDQV GAGGFEFGDKPKITAEAGWVNTAIQGMFVLGLPYAILHGGYGLFLIIIAAVVCC YTKIILLACLYENEDGVVRDYSVAIANACCAFRFTLGRVNVQAILIELVWC ILYVVSGLMYSFPLGVPQSKNSIIATVALLPCFLNKLKNSFLICLAFV INFLVAYLSRRARDWAEKRFYIDVKFPISIGIIVFSYTSQIFPLSLEGNMQOPS EFCMMNTHIAACVILKGLFALVYLWADETKEVITDNLPGSIRAVVNI FLVAKLL SVPLPFAVLEKSLFQEGRAFPACVGGDGRKLSWGLTLRCALVVFLLIMAYV PHFALLMGLTGLTAGLCLFLLPLSLFLHLLWRKLLWHQVFFDVAIFVIGICSVSGF VHSLEGLIEAYRNAED"
ORIGIN	
Query Match	96.2%; Score 1518.8; DB 9; Length 2756;
Best Local Similarity	97.7%; Pred. No. 1.3e-220;
Matches 1541; Conservative	0; Mismatches 37; Indels 0; Gaps 0;
QY	1 ATGGCCACCTTGCTCGCAGCAAGCTGTCCAAGTGGCCACGCTCGCTGTCCAAAGTCC 60
Db	436 ATGGCCACCTTGCTCGCAGCAAGCTGTCCAAGTGGCCACGCTCGCTGTCCAAAGTCC 495
QY	61 CAGGCCAAGATAGCGCGCATGTTTCGCCAGGATGGGTTTTTCAGCGGCCCGGATGAGGAG 120
Db	496 CAGGCCAAGTGTAGCGGCATGTTTCGCCAGGATGGGTTTTTCAGCGGCCCGGATGAGGAG 555
QY	121 GCGGTGGGCTTCGCGCATTCGACGACCTTCGACTTTGAGCACCACGAGGCGCTGCAATG 180
Db	556 GCGGTGGGCTTCGCGCATTCGACGACCTTCGACTTTGAGCACCACGAGGCGCTGCAATG 615
QY	181 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGAGAGAGCGCTGAGAGCGCGCTCGAG 240
Db	616 GACATCTGAAAGCTGAGGCGAGGCGCTGCGGGAGAGAGGCGCTGAAACGCGCGCTCGAG 675
QY	241 GGAGACATCCATTATCAGCGAGGAGCGAGGCGCTCTTCGCGCGCTCCGCGCTCCAAGGAC 300
Db	676 GGAGACATCCATTATCAGCGCGGAGCGGCGCTCTTCGCGCGCTCCGCGCTCCAAGGAC 735
QY	301 CAGGTGGGAGTGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATACAGGCGTGGAG 360
Db	736 CAGGTGGGAGTGTGGTGAATTCGGGGGCCACGACAAAGCCCAAGATACAGGCGTGGAG 795
QY	361 GCAGGCTGGAACGTGACCAACGCATCCAGGCGATGTTCTGCTGGGCGCTTACCTACGCC 420
Db	796 GCAGGCTGGAACGTGACCAACGCATCCAGGCGATGTTCTGCTGGGCGCTTACCTACGCC 855
QY	421 ATCTCTGACGCGCGCTACCTGGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGCTGC 480
Db	856 ATCTCTGACGCGCGCTACCTGGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGCTGC 915
QY	481 TACACCGGCAAGATCTCATTCGCGTGCCTGTACGAGGAGAAATGAAGACGCGAGGTGGT 540
Db	916 TACACTGGCAAGATCTCATTCGCGTGCCTGTACGAGGAGAAATGAAGACGCGAGGTGGT 975
QY	541 CCGGTGGGAGTCTGTCAGTGGCCATAGCCAAAGCCTGCTGCGCGCGCGCTTCCCAAGC 600
Db	976 CCGGTGGGAGTCTTACGTGGCCATAGCCAAAGCCTGCTGCGCGCGCGCTTCCCAAGC 1035
QY	601 CTGGCGCGCGAGTGTGTGAACGTAGCGCAGATCATCGAGCTGTGTGATGAGCTGCATCTG 660
Db	1036 CTGGCGCGCGCGTGTGTGAACGTGTGGCGAGATCATCGAGCTGTGTGATGAGCTGCATCTG 1095
QY	661 TAGCTGGTGTGAGTGGCAACCTCATGTACAACAGCTTCCCGGGGTGCGCGCTGTCGAG 720
Db	1096 TAGCTGGTGTGAGTGGCAACCTCATGTACAACAGCTTCCCGGGGTGCGCGCTGTCGAG 1155
QY	721 AGTCTGTGTTCATTCGCGCGCGGTGCTGCTGCTGCTGCGCTTCTTATAGAACCTC 780
Db	1156 AGTCTGTGTTCATTCGCGCGCGGTGCTGCTGCTGCTGCGCTTCTTATAGAACCTC 1215
QY	781 AAGGCGCGTGTCCAAGTTCAGTGTGTCACCTCTGGCCCACTTCGTTCATCAATATCCTG 840
Db	1216 AAGGCGCGTGTCCAAGTTCAGTGTGTCACCTCTGGCCCACTTCGTTCATCAATATCCTA 1275
QY	841 GTCATAGCTACTGTCTTATCGGGGGCGCGAGCTGGGCTGGGAGAGGTCGAGTTCTAC 900
Db	1276 GTCATAGCTACTGTCTTATCGGGGGCGCGAGCTGGGCTGGGAGAGGTCGAGTTCTAC 1335
QY	901 ATCGAGCTCAAGAGTTCCCATCTCCATTTGGCATCATGTGTTCAGCTACACGTCTCAG 960
Db	1336 ATCGAGCTCAAGAGTTCCCATCTCCATTTGGCATCATGTGTTCAGCTACACGTCTCAG 1395
QY	961 ATCTTCTGCTTCTGCTGGAGGGCAATATGACAGCAGCCAGCGAGTTCACACTGCATGATG 1020
Db	1396 ATCTTCTGCTTCTGCTGGAGGGCAATATGACAGCAGCCAGCGAGTTCACACTGCATGATG 1455
QY	1021 AACTGGAAGCAGCATCGAGCTGCGGTCTCAAGGGGCTCTTCGCGCTGCTGCGCTACCTC 1080
Db	1456 AACTGGAAGCAGCATCGAGCTGCGGTCTCAAGGGGCTCTTCGCACTCGTGGCTACCTC 1515
QY	1081 ACCTGGGCGCAGAGACCAAGAGGTTCATCAAGGATAACTGCGCGGCTCCATCCGGGC 1140
Db	1516 ACCTGGGCGCAGAGACCAAGAGGTTCATCAAGGATAACTGCGCGGCTCCATCCGGGC 1575
QY	1141 GTGGTCAACATCTTTCTGCTGGCGCAAGGGCGCTGTTCTCTATCTCTGCGCATTTCTTGGC 1200

LOCUS AX497266 1733 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 33 from Patent WO0229058.
ACCESSION AX497266
VERSION AX497266.1 GI:23342642

KEYWORDS

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Shinketsu, R.A., Taupier, R.J., Burgess, C.E., Zerhusen, B.D.,
Mezes, P.S., Rastelli, L., Malyankar, U.M., Grosse, W.M.,
Alsbrook, J.P., Lepley, D.M., Spytek, K.A., Li, L., Edinger, S.,
Gerlach, V., Ellemman, K., Macdougall, J., Gunther, E., Millet, I.,
Stone, D., Smithson, G., and Szekeres, E.S.
TITLE Human proteins, polynucleotides encoding them and methods of using
the same

JOURNAL Patent: WO 0229058-A 33 11-APR-2002;

Curagen Corporation (US)

FEATURES

source
1..1733
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.3%; Score 1567.4; DB 6; Length 1733;
Best Local Similarity 99.9%; Pred. No. 6.2e-228;
Matches 1568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTCTCGCAGCAAGCTGTCCAACTGGCCACGCTCGGTGTCCAACTGTC 60
DB 112 ATGGCCACCTTCTCGCAGCAAGCTGTCCAACTGGCCACGCTCGGTGTCCAACTGTC 171
QY 61 CAGGCCAAGATGAGCGGATGTTTCGCGAGGATGGGTTTTCAGCGGCGCAGGATGAGGAG 120
DB 172 CAGGCCAAGATGAGCGGATGTTTCGCGAGGATGGGTTTTCAGCGGCGCAGGATGAGGAG 231
QY 121 GCGGTGGGCTTCGCGATTTGCGACGACCTCGACTTTTTCAGCAGCGCCGCTGAGATG 180
DB 232 GCGGTGGGCTTCGCGATTTGCGACGACCTCGACTTTTTCAGCAGCGCCGCTGAGATG 291
QY 181 GACATCTCTGAAACCGGAGGAGCGCTTCGCGGAGCAGGCGCTGAAGCGCCGCTCGAG 240
DB 292 GACATCTCTGAAACCGGAGGAGCGCTTCGCGGAGCAGGCGCTGAAGCGCCGCTCGAG 351
QY 241 GGAGACATCCATTATCAGCGAGCGAGCGAGCTCTCTCGCGCCCTCGGCTCCAGGAC 300
DB 352 GGAGACATCCATTATCAGCGAGCGAGCGAGCTCTCTCGCGCCCTCGGCTCCAGGAC 411
QY 301 CAGGTGGGAGGTGGTGGCGAATTCGGGGCCACGACCAAGCCAAATCAGCGCGTGGGAG 360
DB 412 CAGGTGGGAGGTGGTGGCGAATTCGGGGCCACGACCAAGCCAAATCAGCGCGTGGGAG 471
QY 361 GCAGGCTGGAACGTGACCAACGCAATTCAGGGCATGTTCTGTCGCGCCCTACCTACGCC 420
DB 472 GCAGGCTGGAACGTGACCAACGCAATTCAGGGCATGTTCTGTCGCGCCCTACCTACGCC 531
QY 421 ATCTCGACGGCGGCTTACCTGGGTTGTTTCTCATCATCTTCGCGCCCTTGTGTCTGC 480
DB 532 ATCTCGACGGCGGCTTACCTGGGTTGTTTCTCATCATCTTCGCGCCCTTGTGTCTGC 591
QY 481 TACACCGGCAAGATCCTCATCGGTCCTGTACGAGGAATGAGACGCGGAGGTGGTG 540
DB 592 TACACCGGCAAGATCCTCATCGGTCCTGTACGAGGAATGAGACGCGGAGGTGGTG 651
QY 541 CCGGTGCGGACTCTGTACGTGGCCATAGCAGCCTGTGCGCCCGCCGCTTCCCAACG 600
DB 652 CCGGTGCGGACTCTGTACGTGGCCATAGCAGCCTGTGCGCCCGCCGCTTCCCAACG 711
QY 601 CTGGGCGGCGAGTGTGTAACGTGAGCGAGATCATCGAGCTGTGTATGATGATGATCTCTG 660
DB 712 CTGGGCGGCGAGTGTGTAACGTGAGCGAGATCATCGAGCTGTGTATGATGATGATGATCTCTG 771

RESULT 13

AB062931

LOCUS

DEFINITION

Macaca fascicularis

sequence.

AB062931

VERSION

2756 bp mRNA linear PRI 13-JUN-2001

Macaca fascicularis brain cDNA clone: QccE-21148, full insert

GI:14388325

AUTHORS

Shimkets, R.A., Taupier, R.J., Burgess, C.E., Zerkusen, B.D.,
Mezes, P.S., Rastelli, L., Malyankar, U.M., Grosse, W.M.,
Alsebrook, J.P., Lepley, D.M., Spytek, K.A., Li, L., Edinger, S.,
Gerlach, V., Ellerman, K., Macdougall, J., Gunther, E., Millet, I.,
Stone, D., Smithson, G. and Szekeres, E.S.
Human proteins, polynucleotides encoding them and methods of using
the same

JOURNAL

Patent: WO 0229058-A 39 11-APR-2002;
Curagen Corporation (US)

FEATURES

source
1. .1734
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

CDS

112. .1689
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD48688.1"
/db_xref="GI:23342648"
/translation="MATLLRSKLSNVATSVNSKSOAKSGMFARMGFOANTDEAVGF
AHCDDLFERHQGMQDILKAEGPCDEGAEPVEGDIHQKSGAPLPPSGKDV
GGGEFDHDKPKITAWERAGWNTNAIQGMFVLGFLYAILHGGYLGFLIIFAAVCC
YTKILLIACIYEENEDGVVRDYSVAIANACAPRFTLGGRVNVAQIIEIWMTC
ILYVVSNGLMYNSFPGLPYVSKWSIIATAVLLPCAPFLKRLKAVSFSLGCTLAHFV
INILVIAVCLSRARDMAWEKVFYIDVKEPISITGIIVFSYTSQILPLPSLEGNMOOPS
EFCMMWTHIAACVLKGLFALVAYLTWADETKEAIDNLPISRIVAVNIFLVAKALL
SYPLPFAAVEVLEKSLFQBSRAFFPACYSGGRLKWSGLTLCALVVFLLMAIIV
PHFALLMGLTSGLTAGLCFLPLSLFHLRLLRLLWHQVFFDVAIFVIGGICSVSGF
VHSLGLEIIEAYRTNAED"

ORIGIN

Query Match 99.8%; Score 1574.8; DB 6; Length 1734;
Best Local Similarity 99.9%; Pred. No. 4.7e-229;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCCACCTTGTCCGAGCAAGCTGTCCAACTGGCCACGTCCTGGTCCAAAGTCC 60
DB 112 ATGCCACCTTGTCCGAGCAAGCTGTCCAACTGGCCACGTCCTGGTCCAAAGTCC 171

QY 61 CAGGCCAAGATAGCGGCATGTTCCGCCAGATGGGTTTTCAGCGCGCCACGATGAGGAG 120
DB 172 CAGGCCAAGATAGCGGCATGTTCCGCCAGATGGGTTTTCAGCGCGCCACGATGAGGAG 231

QY 121 GCGGTGGGCTTCGGCATTCGGACGACCTCGACTTTGAGCACCGCCAGGCGCTGCAGATG 180
DB 232 GCGGTGGGCTTCGGCATTCGGACGACCTCGACTTTGAGCACCGCCAGGCGCTGCAGATG 291

QY 181 GACATCTGAAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 240
DB 292 GACATCTGAAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 351

QY 241 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTGCGGCCCTCGGCTCCAAAGGAC 300
DB 352 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTGCGGCCCTCGGCTCCAAAGGAC 411

QY 301 CAGGTGGGAGTGTGGCGAATTCGGGGGACACAGAACGCCAAATACAGGCGTGGGAG 360
DB 412 CAGGTGGGAGTGTGGCGAATTCGGGGGACACAGAACGCCAAATACAGGCGTGGGAG 471

QY 361 GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGTGTCGGGCTACCTACGCC 420
DB 472 GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGTGTCGGGCTACCTACGCC 531

QY 421 ATCTGCACGCGGCTACCTACCTGGGTTGTTTCTCATCATCTTTCGCGCGCTGTGTCTGC 480
DB 532 ATCTGCACGCGGCTACCTACCTGGGTTGTTTCTCATCATCTTTCGCGCGCTGTGTCTGC 591

QY 481 TACACCGCAAGATCTCATCGGTGCTGTACAGGAGATGAAGACGCGGAGTGGTG 540
DB 592 TACACCGCAAGATCTCATCGGTGCTGTACAGGAGATGAAGACGCGGAGTGGTG 651

QY 541 CGCGTGGGACTCGTACGTGGCCATAGCCAAAGCTGTGCGGCCCGCGCTTCCCAACG 600

RESULT 12
AX497266

source	1. .1734	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
CDS	112. .1689	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD48687.1"	
	/db_xref="GI:23342646"	
	/translation="WATLLRSKLSNVATSVSNKSQAKSMGFARMFQOATDEEAVGF AHCDDLEFHRQGLQMDILKAEFGPCGDEAEPVEGDIHYORGSGAPLPSPGSKDQV GGGFGHGDKPKITAEWAGMVTNAIQGMFVLGYPYAILHGGYLGLEPLIPAAVCC YTKILLIACLYENEDGEVVRDYSVAIANACCAPRPTLGGRRVNVQAIIIEIWMTC ILYVNSGLMYNSFPGLPVSKOSIIATAPLPCAFKLNKAVSKFSLLEGTLLAHFV INILVAYCLSRDWDWAEKVFYIDVKKFPISIGIIVFSYTSQIFLPSLEGNMQOPS EPEFMNWHIAACVLKGLPALVALVLTWADETKAITDNLPGSIRAVNVIPLVAKLL SYPLPFAAIVELEKSLFQBSRAFFPACYSGDRKLSWGLTLRLCALVFLLMAYIV PFPALLMGLTSLTGAGLCFLPLPSLPHLLWRKLLMHQVFPDVAIFVIGGICVSGF VHSLEGLIEAYRTNAED"	
ORIGIN		
Query Match	99.8%;	Score 1574.8; DB 6; Length 1734;
Best Local Similarity	99.9%;	Pred. No. 4.7e-229;
Matches 1576;	Conservative 0;	Mismatches 2; Indels 0; Gaps 0;
QY	1	ATGGCCACCTTGCTCGCAGCAAGCTGTCCAACTGGCCACGCTCGTGTCCAAAGTCC 60
DB	112	ATGGCCACCTTGCTCGCAGCAAGCTGTCCAACTGGCCACGCTCGTGTCCAAAGTCC 171
QY	61	CAGGCCAAGATGAGCGGCATGTCGCGAGATGGGTTTTCAGCGCGCCACGATGAGGAG 120
DB	172	CAGGCCAAGATGAGCGGCATGTCGCGAGATGGGTTTTCAGCGCGCCACGATGAGGAG 231
QY	121	GGGTGGGCTTCGCGCATTCGCGACGACCTCGACTTTGACGACCGCGAGGCGCTGCAGATG 180
DB	232	GGGTGGGCTTCGCGCATTCGCGACGACCTCGACTTTGACGACCGCGAGGCGCTGCAGATG 291
QY	181	GACATCTCAAGACCGAGGAGAGCCCTCGCGGAGCGAGGCGGCTGAAGCGCCCTCGAG 240
DB	292	GACATCTCAAGACCGAGGAGAGCCCTCGCGGAGCGAGGCGGCTGAAGCGCCCTCGAG 351
QY	241	GGAGACATCCATATATCAGCGAGGAGCGGAGCTCTCTCGCGCCCTCGGCTCCAAAGGAC 300
DB	352	GGAGACATCCATATATCAGCGAGGAGCGGAGCTCTCTCGCGCCCTCGGCTCCAAAGGAC 411
QY	301	CAGGTGGAGAGTGGTGGCAATTCGGGGGCGACGACCAAGCCCAAAATCAAGCGGTGGAG 360
DB	412	CAGGTGGAGAGTGGTGGCAATTCGGGGGCGACGACCAAGCCCAAAATTAAGCGGTGGAG 471
QY	361	GCAGGCTGGAACTGACCAACGCGCATCCAGGGCATGTTGTTGGGCGCTTACCCCTACGCC 420
DB	472	GCAGGCTGGAACTGACCAACGCGCATCCAGGGCATGTTGTTGGGCGCTTACCCCTACGCC 531
QY	421	ATCCTGCACGCGCGCTACCTGGGTTGTTTCTCATCATCTTCGCGCCGCTTGTGTCTGC 480
DB	532	ATCCTGCACGCGCGCTACCTGGGTTGTTTCTCATCATCTTCGCGCCGCTTGTGTCTGC 591
QY	481	TACACCGCAAGATCCTCATTCGGTGGCTGTACGAGGAGAAATGAAGACGGGAGTGGTG 540
DB	592	TACACCGCAAGATCCTCATTCGGTGGCTGTACGAGGAGAAATGAAGACGGGAGTGGTG 651
QY	541	CGGTCGGGAGCTGTAGTGGCCATAGCCCAAGCTGCTGCGCCCGCTTCCCAAGC 600
DB	652	CGGTCGGGAGCTGTAGTGGCCATAGCCCAAGCTGCTGCGCCCGCTTCCCAAGC 711
QY	601	CTGGGCGGCGGAGTGGTGAACGCTAGCGGAGATCATCGAGTGGTGAATGACGTGCACTCTG 660
DB	712	CTGGGCGGCGGAGTGGTGAACGCTAGCGGAGATCATCGAGTGGTGAATGACGTGCACTCTG 771
QY	661	TACGTGGTGGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCCCGGTGTCGAG 720
DB	772	TACGTGGTGGTGGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCCCGGTGTCGAG 831
RESULT 11		
AX497272		
Sequence 39 from Patent W00229058.		
AX497272		
AX497272.1 GI:23342647		
Homo sapiens (human)		
Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE 1		
PAT 26-SEP-2002		

AHCDLDFEHRQQLQMDILKAEGPCDEGAEPVEGDHYQRSGAPLPPSSKQV
GGGEGFGHDKPKITAEAGVNVNATQMFVLGYPVAILHGGYGLFLFIIFAAVCC
YTKGILILAEYENEDGVVRDVAIAIANACAPFPTLGGRVNVAOIIELVNCC
ILVNVSGNLMYNSFPLPVOKSWSIIATAVLLPCAFLOKXAKVESLCTLAHFV
INILVATCLSRARDNAWEKVIYDVKKFPISIGIIVFYTSQIFLPSLEGNNQOPS
EFHGMNTHIAACVLKGLFALVAYLTKWDEKAITDNLPUSIRAVNIFLVAKALL
SYPLPFAAIVELEKSLFQBSRAFPACYSGDRLKSLWGLTLRICALVWFTLLMAYV
PHFALLMGLTGSITGAGLCFLPLSLFHLRLWRKLLWHQVFDVAIFVIGIGCSVSGF
VHSLEGLIEAYRTNAED"

ORIGIN

Query Match 99.8%; Score 1574.8; DB 6; Length 1734;
Best Local Similarity 99.9%; Pred. No. 4.7e-229;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGCCACCTTGTCCGACAGCTGTCCAAAGTGGCCACGTCCTGGTCCAAAGTCC	60
Db	112	ATGGCCACCTTGTCCGACAGCTGTCCAAAGTGGCCACGTCCTGGTCCAAAGTCC	171
QY	61	CAGGCCAAGATGAGCGGCATGTTCCGACAGATGGTTTTCAGGCGGCCACGGATGAGGAG	120
Db	172	CAGGCCAAGATGAGCGGCATGTTCCGACAGATGGTTTTCAGGCGGCCACGGATGAGGAG	231
QY	121	GGGTGGGCTTCGGCGATTCGACGACCTCGACTTTTGAGCAACGCCAGGCGCTTCGACATG	180
Db	232	GGGTGGGCTTCGGCGATTCGACGACCTCGACTTTTGAGCAACGCCAGGCGCTTCGACATG	291
QY	181	GACATCTCTGAAGCCGAGGAGAGCCCTCGCGGAGCGAGGCGCTGAAGCGCCGCTCGAG	240
Db	292	GACATCTCTGAAGCCGAGGAGAGCCCTCGCGGAGCGAGGCGCTGAAGCGCCGCTCGAG	351
QY	241	GGAGACATCATTTATCAGCGAGGAGCGGAGCTCTCTCGCGCCCTCCGGCTCCAAAGGAC	300
Db	352	GGAGACATCATTTATCAGCGAGGAGCGGAGCTCTCTCGCGCCCTCCGGCTCCAAAGGAC	411
QY	301	CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCACGGCGTGGGAG	360
Db	412	CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCACGGCGTGGGAG	471
QY	361	GCAGGTGGAGAGCTGACCAACGCGATCCAGGAGATGTTCTGTCTGGGCTTACCTACGCC	420
Db	472	GCAGGTGGAGAGCTGACCAACGCGATCCAGGAGATGTTCTGTCTGGGCTTACCTACGCC	531
QY	421	ATCTGCACGCGCGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGCTGTGCTGC	480
Db	532	ATCTGCACGCGCGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGCTGTGCTGC	591
QY	481	TACACCGGCAAGATCTCTCATCGCGTGCCTGTACGAGGAGAAATGAAGCGCGAGGTGGTG	540
Db	592	TACACCGGCAAGATCTCTCATCGCGTGCCTGTACGAGGAGAAATGAAGCGCGAGGTGGTG	651
QY	541	CGGCTGGGAGCTGTACGTGGCGATAGCAACGCTGTGCGCGCGCGCTTCCCAACG	600
Db	652	CGGCTGGGAGCTGTACGTGGCGATAGCAACGCTGTGCGCGCGCGCTTCCCAACG	711
QY	601	CTGGCGCGCGAGTGGTGAAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT	660
Db	712	CTGGCGCGCGAGTGGTGAAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT	771
QY	661	TACGTGGTGGTGAAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCGCGCTGCGCAG	720
Db	772	TACGTGGTGGTGAAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCGCGCTGCGCAG	831
QY	721	AAGTCTGTGTCATATATCGCACGCGCGTGTCTGCTGCTTGGCGCTTCTTAAAGACCTC	780
Db	832	AAGTCTGTGTCATATATCGCACGCGCGTGTCTGCTGCTTGGCGCTTCTTAAAGACCTC	891
QY	781	AAGCGGTGTCGAAGTTTCACTGTGCTGTGACCTCTGCGCCACTTCGTTCATATATCTC	840
Db	892	AAGCGGTGTCGAAGTTTCACTGTGCTGTGACCTCTGCGCCACTTCGTTCATATATCTC	951
QY	841	GTCCATAGCTTACTGTCTATCGCGGCGCGGACTGGGCGCTGGGAGAGGTCAAGTTCTAC	900

Db	952	GTCCATAGCTTACTGTCTATCGCGGCGCGGACTGGGCGCTGGGAGAGGTCAAGTTCTAC	1011
QY	901	ATCGAGCTCAAGAAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGTCTCAG	960
Db	1012	ATCGAGCTCAAGAAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGTCTCAG	1071
QY	961	ATCTTCTCGCTTTCGCTGGAGGGCAATATGACAGCAGCCAGCGAGTTCACATGATGATG	1020
Db	1072	ATCTTCTCGCTTTCGCTGGAGGGCAATATGACAGCAGCCAGCGAGTTCACATGATGATG	1131
QY	1021	AACTGAGCAGCATATCGAGCTCGCTCAAGGGCTCTTTCGCGCTGTGCGCTACCTC	1080
Db	1132	AACTGAGCAGCATATCGAGCTCGCTCAAGGGCTCTTTCGCGCTGTGCGCTACCTC	1191
QY	1081	ACCTGGGCGCAGCAGACCAAGAGGTATACACGATAACTGCGCGGCTCCATCCGCGCC	1140
Db	1192	ACCTGGGCGCAGCAGACCAAGAGGTATACACGATAACTGCGCGGCTCCATCCGCGCC	1251
QY	1141	GTGGTCAACATCTTCTGTGGTGGCAAGGGCTCTTGTCTTATCTCTGCGCATTTCTTGGC	1200
Db	1252	GTGGTCAACATCTTCTGTGGTGGCAAGGGCTCTTGTCTTATCTCTGCGCATTTCTTGGC	1311
QY	1201	GCTGTGAGGTGCTGAGAAAGTCTTCCAGAAAGGAGCAGCGCGCTTTTTCGCGGCC	1260
Db	1312	GCTGTGAGGTGCTGAGAAAGTCTTCCAGAAAGGAGCAGCGCGCTTTTTCGCGGCC	1371
QY	1261	TGCTACAGCGCGCAGCGGCGCTTGAAGTCTGTGGGGCTGACGCTGGGCTGGCGCTGCTC	1320
Db	1372	TGCTACAGCGCGCAGCGGCGCTTGAAGTCTGTGGGGCTGACGCTGGGCTGGCGCTGCTC	1431
QY	1321	GTCTTACGCTGCTCATGGCCATTTATGTGCCGCACTTGGCGCTGTCTATGGGCTTCACC	1380
Db	1432	GTCTTACGCTGCTCATGGCCATTTATGTGCCGCACTTGGCGCTGTCTATGGGCTTCACC	1491
QY	1381	GGCAGCTCAAGCGCGCGCTTCTGTTTCTGCTGCCAGCTCTTTCACCTGCGCTG	1440
Db	1492	GGCAGCTCAAGCGCGCGCTTCTGTTTCTGCTGCCAGCTCTTTCACCTGCGCTG	1551
QY	1441	CTTGGGCGCAGCTGTGTGGCAACCAAGTCTTTCGAGTGGCGCTTCTGCTATCGTCCG	1500
Db	1552	CTTGGGCGCAGCTGTGTGGCAACCAAGTCTTTCGAGTGGCGCTTCTGCTATCGTCCG	1611
QY	1501	GGCATCTGAGCTGCTCGGCTTCTGTCGCTCCTCGAGGCGCTCATCGAAGCTTACCGA	1560
Db	1612	GGCATCTGAGCTGCTCGGCTTCTGTCGCTCCTCGAGGCGCTCATCGAAGCTTACCGA	1671
QY	1561	ACCAACGCGGAGGACTAG 1578	
Db	1672	ACCAACGCGGAGGACTAG 1689	

RESULT 10

AX497270	LOCUS	1734 bp	DNA	linear	PAT 26-SEP-2002
AX497270	DEFINITION	Sequence 37 from Patent WO0229058.			
AX497270	ACCESSION	AX497270			
AX497270.1	VERSION	GI:23342645			
	KEYWORDS				
	SOURCE	Homo sapiens (human)			
	ORGANISM	Homo sapiens			
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	REFERENCE				
	AUTHORS	Shinkets, R.A., Taupier, R.J., Burgess, C.E., Zerhusen, B.D.,			
		Mezes, P.S., Rastelli, L., Malyankar, U.M., Grosse, W.M.,			
		Alsobrook, J.P., Lepley, D.M., Spytek, K.A., Li, L., Edinger, S.,			
		Gerlach, V., Ellerman, K., Macdougall, J., Gunther, E., Millet, I.,			
		Stone, D., Ellerman, G. and Szekeres, E.S.			
	TITLE	Human proteins, polynucleotides encoding them and methods of using			
		the same			
	JOURNAL	Patent: WO 0229058-A 37 11-APR-2002;			
		Curagen Corporation (US)			
	FEATURES	Location/Qualifiers			

Query Match 99.8%; Score 1574.8; DB 6; Length 1734;
Best Local Similarity 99.9%; Pred No. 4.7e-229;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTCTCGCAGCAAGCTGTCAAAGTGGCCATCTCGTGTCAAACAGTCC 60
DB 112 ATGGCCACCTTCTCGCAGCAAGCTGTCAAAGTGGCCATCTCGTGTCAAACAGTCC 171

QY 61 CAGGCCAGATCAGCGGCATGTTCCAGGATGGTTTTCAGCGCCGACGATCAGGAG 120
DB 172 CAGGCCAGATCAGCGGCATGTTCCAGGATGGTTTTCAGCGCCGACGATCAGGAG 231

QY 121 CGCGTGGGCTTCGCGCATTTGCGACACCTTCGACTTTTGAACCGCAGCGCCCTGCAGATG 180
DB 232 CGCGTGGGCTTCGCGCATTTGCGACACCTTCGACTTTTGAACCGCAGCGCCCTGCAGATG 291

QY 181 GACATCTGAAGCCGAGGAGAGCCCTCGCGGAGAGGGCGCTGGAAGCGCCGCTCGAG 240
DB 292 GACATCTGAAGCCGAGGAGAGCCCTCGCGGAGAGGGCGCTGGAAGCGCCGCTCGAG 351

QY 241 GGAGACATCATATCAGCGGACGAGGAGCTCTCTCGCGCCCTCGGCTCCAGGAC 300
DB 352 GGAGACATCATATCAGCGGACGAGGAGCTCTCTCGCGCCCTCGGCTCCAGGAC 411

QY 301 CAGGTGGGAGGTGGTGGCAATTCGGGGGCCACGACCAAGCCCAAAATCACGGCGTGGAG 360
DB 412 CAGGTGGGAGGTGGTGGCAATTCGGGGGCCACGACCAAGCCCAAAATCACGGCGTGGAG 471

QY 361 CGAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGCTCGGCGCTACCTACGCC 420
DB 472 CGAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGCTCGGCGCTACCTACGCC 531

QY 421 ATCTCTGACGGCGGTACCTGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGCTGC 480
DB 532 ATCTCTGACGGCGGTACCTGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGCTGC 591

QY 481 TACACCGCAAGATCTCATCGCTGCTGTAGGAGGAATGAAGACGGCGAGGTGGTG 540
DB 592 TACACCGCAAGATCTCATCGCTGCTGTAGGAGGAATGAAGACGGCGAGGTGGTG 651

QY 541 CGCGTGGGACTCGTACGTGGCCATGACCAAGCTGTGCGCCCGCGCTTCCCAACG 600
DB 652 CGCGTGGGACTCGTACGTGGCCATGACCAAGCTGTGCGCCCGCGCTTCCCAACG 711

QY 601 CTGGCGCGCGAGTGTGAAGCTAGCGGACATCATCGAGCTGTGATGACGTGCATCTTG 660
DB 712 CTGGCGCGCGAGTGTGAAGCTAGCGGACATCATCGAGCTGTGATGACGTGCATCTTG 771

QY 661 TACGTGTGTGAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCGCGTGTGCGAG 720
DB 772 TACGTGTGTGAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCGCGTGTGCGAG 831

QY 721 AAGTCTGTGTCCATATCGCCAGCGCGTGTGCTGCTGCTGCGCTTCCCTTAAGAACCTTC 780
DB 832 AAGTCTGTGTCCATATCGCCAGCGCGTGTGCTGCTGCTGCGCTTCCCTTAAGAACCTTC 891

QY 781 AAGCGCGTGTCAAAGTTGAGTGTGCACTCTGCGCCCACTTCTGTCATCAATATCTTG 840
DB 892 AAGCGCGTGTCAAAGTTGAGTGTGCACTCTGCGCCCACTTCTGTCATCAATATCTTG 951

QY 841 GTCATAGCTTACTGTCTATCGCGGGCGCGACTGGGCTGGGAGAGGTCAAGTTCTAC 900
DB 952 GTCATAGCTTACTGTCTATCGCGGGCGCGACTGGGCTGGGAGAGGTCAAGTTCTAC 1011

QY 901 ATCGAGCTCAAGAAGTTCCCATCTCCATTTGGCATCATGTTGTTCAGCTACAGTCTCAG 960
DB 1012 ATCGAGCTCAAGAAGTTCCCATCTCCATTTGGCATCATGTTGTTCAGCTACAGTCTCAG 1071

QY 961 ATCTTCTGCTTTCGTGGAGGGCAATATGACAGCCAGCCAGCGAGTTCACCTGATGATG 1020
DB 1072 ATCTTCTGCTTTCGTGGAGGGCAATATGACAGCCAGCCAGCGAGTTCACCTGATGATG 1131

RESULT 9
AX497268
LOCUS AX497268 1734 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 35 from Patent WO0229058.
ACCESSION AX497268
VERSION AX497268.1 GI:23342643
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Shinkets, R.A., Taupier, R.J., Burgess, C.E., Zerhusen, B.D.,
Mezes, P., Rastelli, L., Malyankar, U.M., Grosse, W.M.,
Alsobrook, J.P., Lepley, D.M., Spytek, K.A., Li, L., Edinger, S.,
Gerlach, V., Ellerman, K., Macdougall, J., Gunther, E., Millet, I.,
Stone, B., Smithson, G. and Szekeres, E.S.
Human proteins, polynucleotides encoding them and methods of using
the same
Patent: WO 0229058-A 35 11-APR-2002;
Curagen Corporation (US)
Location/Qualifiers
1..1734
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
112..1689
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD48686.1"
/db_xref="GI:23342644"
/translation="MATLLRSKLSNVATSVSNKSQAKMSGMPFARMGFOATDEAVGF"

TITLE
JOURNAL
FEATURES
source
CDS

Db 261 GCGGTGGGCTTCGCGATTCGACGACCTCGACTTTGAGCACCCGCGCTCGAGATG 320
Qy 181 GACATCTTAAAGCCGAGGAGAGCCCTGCGGGACGAGGCGCTGAAGCCCGCTCGAG 240
Db 321 GACATCTTAAAGCCGAGGAGAGCCCTGCGGGACGAGGCGCTGAAGCCCGCTCGAG 380
Qy 241 GAGAGACATTCATTTATCAGCGAGCGAGCGAGCTCTCTGCGGCCCTCGGGCTCAAGGAC 300
Db 381 GAGAGACATTCATTTATCAGCGAGCGAGCGAGCTCTCTGCGGCCCTCGGGCTCAAGGAC 440
Qy 301 CAGGTGGAGGTGGTGGCAATTCGGGGCCACGACAGCCCAAAATCAGCGCGTGGAG 360
Db 441 CAGGTGGAGGTGGTGGCAATTCGGGGCCACGACAGCCCAAAATCAGCGCGTGGAG 500
Qy 361 CAGGCTGGAGCGTGAACCAAGCCATCCAGGCGATGTCGTCTGGGCTACCTACGCC 420
Db 501 CAGGCTGGAGCGTGAACCAAGCCATCCAGGCGATGTCGTCTGGGCTACCTACGCC 560
Qy 421 ATCTGTCACGGCGGCTACCTGGGGTTGTTTCTCATCATCTTCGCGCGGTTGTGTGC 480
Db 561 ATCTGTCACGGCGGCTACCTGGGGTTGTTTCTCATCATCTTCGCGCGGTTGTGTGC 620
Qy 481 TACACGGCAAGATCTCATCGCTGCTGATGACGAGGAATGAAGACGGCGAGTGTG 540
Db 621 TACACGGCAAGATCTCATCGCTGCTGATGACGAGGAATGAAGACGGCGAGTGTG 680
Qy 541 CCGGTGGGAGCTGCTGAGTGGCCATAGCAAGCGCTGCTGCGCCCGCTTCCCAAG 600
Db 681 CCGGTGGGAGCTGCTGAGTGGCCATAGCAAGCGCTGCTGCGCCCGCTTCCCAAG 740
Qy 601 CTGGGCGGCGAGTGTGAACCTAGCGAGATCATCGAGCTGGTGAACGATCTCTG 660
Db 741 CTGGGCGGCGAGTGTGAACCTAGCGAGATCATCGAGCTGGTGAACGATCTCTG 800
Qy 661 TACGTGGTGGTGAAGTGAACCTATGTAACAGCTTCCGGGCTGCTGTCGAG 720
Db 801 TACGTGGTGGTGAAGTGAACCTATGTAACAGCTTCCGGGCTGCTGTCGAG 860
Qy 721 AGTCTGTGTCATTCAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 861 AGTCTGTGTCATTCAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 920
Qy 781 AAGGCGGCTGCAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 921 AAGGCGGCTGCAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 980
Qy 841 GTATAGCTTACTGTCTATCGGGGCGCGAGCTGGGCTGGAGAGGTCAAGTTCATC 900
Db 981 GTATAGCTTACTGTCTATCGGGGCGCGAGCTGGGCTGGAGAGGTCAAGTTCATC 1040
Qy 901 ATCGAGCTCAAGAGTTCCTCATCTCCATTCGATGCTGCTGCTGCTGCTGCTGCTG 960
Db 1041 ATCGAGCTCAAGAGTTCCTCATCTCCATTCGATGCTGCTGCTGCTGCTGCTGCTG 1100
Qy 961 ATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 1101 ATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1160
Qy 1021 AACTGAGCAGATTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1161 AACTGAGCAGATTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1220
Qy 1081 ACTTGGGCGCAGAGAGGCTCATCAGGATAACTGCGCGCTGCTGCTGCTGCTGCTG 1140
Db 1221 ACTTGGGCGCAGAGAGGCTCATCAGGATAACTGCGCGCTGCTGCTGCTGCTGCTG 1280
Qy 1141 GTGGTCAACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 1281 GTGGTCAACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1340
Qy 1201 GCTGTGAGGTGCTGAGAGTCTCTTCCAGAGAGGCGCGCTGCTGCTGCTGCTGCTG 1260
Db 1341 GCTGTGAGGTGCTGAGAGTCTCTTCCAGAGAGGCGCGCTGCTGCTGCTGCTGCTG 1400

Qy 1261 TGCTACAGCGCGAGCGCGCTGCAAGTCTCTGGGGCTGACGCTCGCTGCGCTCGTC 1320
Db 1401 TGCTACAGCGCGAGCGCGCTGCAAGTCTCTGGGGCTGACGCTCGCTGCGCTCGTC 1460
Qy 1321 GTCTTACAGCTGCTCATGCGCATTTATGTGCGGCACTTTCGCGCTGCTCATGGGCTCAC 1380
Db 1461 GTCTTACAGCTGCTCATGCGCATTTATGTGCGGCACTTTCGCGCTGCTCATGGGCTCAC 1520
Qy 1381 GGCAGCTTACAGCGCGCGCGCTGTTCTTCTGCTGCCAGCCTTTTTCACCTGCGCCTG 1440
Db 1521 GGCAGCTTACAGCGCGCGCGCTGTTCTTCTGCTGCCAGCCTTTTTCACCTGCGCCTG 1580
Qy 1441 CTCTGGCCAAAGCTGCTGTGGCACAAGTCTTCTTCGACGCTCGGCATCTTCTGTCATCGGC 1500
Db 1581 CTCTGGCCAAAGCTGCTGTGGCACAAGTCTTCTTCGACGCTCGGCATCTTCTGTCATCGGC 1640
Qy 1501 GGCATCTGACGCTGCTCGGCTTCTGTCATCTTCTGAGGGCTCATCGAGCTTACCGA 1560
Db 1641 GGCATCTGACGCTGCTCGGCTTCTGTCATCTTCTGAGGGCTCATCGAGCTTACCGA 1700
Qy 1561 ACCAAGCGGAGGACTAG 1578
Db 1701 ACCAAGCGGAGGACTAG 1718

RESULT 7
AX259431
LOCUS AX259431 1800 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 1 from Patent WO0173015.
ACCESSION AX259431
VERSION AX259431.1 GI:16508618
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
Rippmann, F. and Duecker, K.
Identification of human gaba transporter
Patent: WO 0173015-A 1 04-OCT-2001;
MERCK PATENT GmbH (DE)
FEATURES
Location/Qualifiers
1..1800
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
100..1677
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10188.1"
/db_xref="GI:16508618"

CDS
1..1800
/translation="MATLLRSLKSLNATSVSNKSKMGMPARMGFQATDEAVGF
AHCDDLDPEHROGLQMDILKAEPCGDEGAEPVEGDHYQRGSGAPLPSPGSKDQV
GGGEGFGHDKPKITANEAGWNTNIAIGFVGLPFIILHGGYGLGLFIIPAAVVC
VTGKLLIACLYEENEDGVRVSYVAIANACCAAPRPTLGGRVVYVAQIITELVMT
ILYVVSGLNMYNSPFGIPVSKWSIIATAVLLPCAFKLNKLVKSKSLCTLAHFV
INILVIACTLRARDWAKWPKYIDVKKFPISIGIIVFSYQIPLPSLEGNMQPS
EFHGMNTHIAACVLKGFALVAYLTWADEKTEVITNLPGRIRAVNIFLVAKLL
SYLPFFPAVEVLEKSLFQEGSRAPFPACYSGDGRKSLWGLTLRCLVALVFTLLMAIYV
PHFALLMGLTGLTAGLCLFLLPSLFLRLMLRHQVFFDVAIFVIGGICSVSGF
VHSLGLELIEAVRTNAED"

ORIGIN
Query Match 99.9%; Score 1576.4; DB 6; Length 1800;
Best Local Similarity 99.9%; Pred. No. 2.7e-229;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGCCACTTGTCTCGCAGCAAGCTGTCCACAGTGGCCACCGCTCGTGTCCAAAGTCC 60
Db 100 ATGGCCACTTGTCTCGCAGCAAGCTGTCCACAGTGGCCACCGCTCGTGTCCAAAGTCC 159
Qy 61 CAGGCCAAGATGAGCGCATGTTCCAGGATGGGTTTTTTCAGCGCGCCACGATGAGGAG 120

QY 1261 TGCTACAGCGGCGAGCGGCGCTGAAGTCTCTGGGGGCTGACGCTGCGCTGCGCGCTCGTC 1320
Db 1435 TGCTACAGCGGCGAGCGGCGCTGAAGTCTCTGGGGGCTGACGCTGCGCTGCGCGCTCGTC 1494
QY 1321 GTCTTACGCTGCTCATGCGCCATTTATGTCGCGCATCTTTCGCGCTGCTCATGCGCGCTCAC 1380
Db 1495 GTCTTACGCTGCTCATGCGCCATTTATGTCGCGCATCTTTCGCGCTGCTCATGCGCGCTCAC 1554
QY 1381 GCGAGCTCACGCGGCGGCGCTGTTCTTCTGCTGCGCGCTCTTTTACCTGCGCGCTG 1440
Db 1555 GCGAGCTCACGCGGCGGCGCTGTTCTTCTGCTGCGCGCTCTTTTACCTGCGCGCTG 1614
QY 1441 CTCTGCGCAAGCTGCTGTGGCAACCAAGTCTTCTTTCGAGCTGCGCATCTTCTGTCATCGGC 1500
Db 1615 CTCTGCGCAAGCTGCTGTGGCAACCAAGTCTTCTTTCGAGCTGCGCATCTTCTGTCATCGGC 1674
QY 1501 GGCATCTGAGGCTGCTGCGGCTTCTGTCATCTTCTGAGGCGCTCATGCAAGCTTACCGA 1560
Db 1675 GGCATCTGAGGCTGCTGCGGCTTCTGTCATCTTCTGAGGCGCTCATGCAAGCTTACCGA 1734
QY 1561 ACCAAGCGGAGGACTAG 1578
Db 1735 ACCAAGCGGAGGACTAG 1752

RESULT 5

AK055051

LOCUS

DEFINITION

AK055051

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT

FEATURES
source

CDS

ORIGIN

Query Match

Best Local Similarity

Matches 1578;

Conservative

0;

Mismatches

0;

Indels

0;

Gaps

0;

Length

2550;

Score

1578;

DB

9;

Pred.

No.

1.4e-229;

100.0%;

100.0%;

100.0%;

100.0%;

100.0%;

100.0%;

100.0%;

100.0%;

1261 TGCTACAGCGGCGAGCGGCGCTGAAGTCTCTGGGGGCTGACGCTGCGCTGCGCGCTCGTC 1320
1435 TGCTACAGCGGCGAGCGGCGCTGAAGTCTCTGGGGGCTGACGCTGCGCTGCGCGCTCGTC 1494
1321 GTCTTACGCTGCTCATGCGCCATTTATGTCGCGCATCTTTCGCGCTGCTCATGCGCGCTCAC 1380
1495 GTCTTACGCTGCTCATGCGCCATTTATGTCGCGCATCTTTCGCGCTGCTCATGCGCGCTCAC 1554
1381 GCGAGCTCACGCGGCGGCGCTGTTCTTCTGCTGCGCGCTCTTTTACCTGCGCGCTG 1440
1555 GCGAGCTCACGCGGCGGCGCTGTTCTTCTGCTGCGCGCTCTTTTACCTGCGCGCTG 1614
1441 CTCTGCGCAAGCTGCTGTGGCAACCAAGTCTTCTTTCGAGCTGCGCATCTTCTGTCATCGGC 1500
1615 CTCTGCGCAAGCTGCTGTGGCAACCAAGTCTTCTTTCGAGCTGCGCATCTTCTGTCATCGGC 1674
1501 GGCATCTGAGGCTGCTGCGGCTTCTGTCATCTTCTGAGGCGCTCATGCAAGCTTACCGA 1560
1675 GGCATCTGAGGCTGCTGCGGCTTCTGTCATCTTCTGAGGCGCTCATGCAAGCTTACCGA 1734
1561 ACCAAGCGGAGGACTAG 1578
1735 ACCAAGCGGAGGACTAG 1752

AK055051 2550 bp mRNA linear PRI 30-JAN-2004
Homo sapiens cDNA FLJ30489 fis, clone BRAWH2000142, highly similar
to Rattus norvegicus vesicular GABA transporter (VGAT) mRNA.
AK055051
AK055051.1 GI:16549699
oligo capping: fis (full insert sequence).
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbe-re@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhgri.nih.gov
 Blakesley, R.W., Bouffard G.G., Green, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 115 Row: c Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 17999519.

FEATURES
source

Location/Qualifiers
 1..2519
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:61677 IMAGE:5180816"
 /tissue_type="Brain, Lung, Testis, adult, pooled whole"
 /clone_lib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 1..2519
 /gene="VIAAT"
 /note="synonym: VGAT"
 /db_xref="LocusID:140679"
 175..1752
 /gene="VIAAT"
 /codon_start=1
 /product="vesicular inhibitory amino acid transporter"
 /protein_id="AAH53582.1"
 /db_xref="GI:31566392"
 /db_xref="LocusID:140679"
 /translations="MATLRSLSNVATSVSNKSQRMGFMARMFOARDDEAVGF
 AHCDLDLDFEHRQGLMDILKAEPEPCDEGAEPVVDGIHYORGSGAPLPPSGSKDQV
 GGGEGFGHDKPXTIWEAGWNVNNAIQGMFVLGLPVAIIHGGYVGLFLIIFAAVVCC
 YTKILLIACLYEENDEGEVVRVDSVAIANACAPRFFPLGRVNVNVAQIIELVNWC
 ILIVVSGNLYNYSFPLPVSVQSKWSIIATAVLLPCAFNLKNAVKFSLLCTLAHFV
 INILVAYCLSRARDWAEVKFYIDVKFPISIGIIVFSYTSQIFLPLEGNMQOPS
 EYFCMMNTHIAACVLKGLFALVAYLTWADETKEVITDNLPGSIRAVVNI FLVAKALL
 SYPLPFAAVLEVKSLFOGSRAPFPACVSGDRLKSWGLTURLCALVFTLIMALVY
 PHEALLGLTGLTGLAGLCFLPLPSLPHLLRLWRKLWHQVFFDVAIFVIGGICSVSGF
 VHSLEGLIEAYRTNAED"

ORIGIN

Query Match 100.0%; Score 1578; DB 9; Length 2519;
 Best Local Similarity 100.0%; Pred. No. 1.4e-229;
 Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ATGGCCACCTTGCTCCGACGAGCTGTCCAAAGTGGCCACGTCGCTGTCCAAAGTCC 60
 175 ATGGCCACCTTGCTCCGACGAGCTGTCCAAAGTGGCCACGTCGCTGTCCAAAGTCC 234
 61 CAGGCCAAGATGAGCGCATGTCGCGAGATCGGTTTCAGGCGGCCGAGGAGGAG 120
 235 CAGGCCAAGATGAGCGCATGTCGCGAGATCGGTTTCAGGCGGCCGAGGAGGAG 294
 121 GCGGTGGGCTTCGCGCATTCGCGAGCTTCGAGCTTTGAGCACCGCCAGGCGCTTCGAGATG 180
 295 GCGGTGGGCTTCGCGCATTCGCGAGCTTCGAGCTTTGAGCACCGCCAGGCGCTTCGAGATG 354

181 GACATCCTGAAAGCCGAGGAGAGCCCTCGGGGAGAGGGCGCTGAAGCGCCGCTCGAG 240
 355 GACATCCTGAAAGCCGAGGAGAGCCCTCGGGGAGAGGGCGCTGAAGCGCCGCTCGAG 414
 241 GGAGACATCATTTATCAGCGAGGAGCGAGGAGCTCTCTGCGGCCCTCCGCGCTCAAGGAC 300
 415 GGAGACATCATTTATCAGCGAGGAGCGAGGAGCTCTCTGCGGCCCTCCGCGCTCAAGGAC 474
 301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCACGGCGTGGGAG 360
 475 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCACGGCGTGGGAG 534
 361 GCGAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGAGGAGCATGTTGGTGGGCTTACCTTAGCC 420
 535 GCGAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGAGGAGCATGTTGGTGGGCTTACCTTAGCC 594
 421 ATCTCTGACGCGCGGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGTGTGCTGC 480
 595 ATCTCTGACGCGCGGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGTGTGCTGC 654
 481 TACACCGGCAAGATCTCTCATTCGCGTCCCTGTACGAGGAGAAATGAAGCGCGAGGTGGTG 540
 655 TACACCGGCAAGATCTCTCATTCGCGTCCCTGTACGAGGAGAAATGAAGCGCGAGGTGGTG 714
 541 CCGGTGGGAGCTCGTACGTGGCCATAGCCAAAGCCCTGCTGCGCGCGCGCTTCCCAAGC 600
 715 CCGGTGGGAGCTCGTACGTGGCCATAGCCAAAGCCCTGCTGCGCGCGCGCTTCCCAAGC 774
 601 CTGGGGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGAGCTGCATCTCTG 660
 775 CTGGGGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGAGCTGCATCTCTG 834
 661 TAGCTGGTGGTGGTGGCAACCTCATGTATCAACAGACTTCCCGGGGCTGCGCGCTGTGCGAG 720
 835 TAGCTGGTGGTGGTGGCAACCTCATGTATCAACAGACTTCCCGGGGCTGCGCGCTGTGCGAG 894
 721 AGTCTCTGGTCCATTTATCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTTCCCTTAGAAGCTC 780
 895 AGTCTCTGGTCCATTTATCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTTCCCTTAGAAGCTC 954
 781 AAGCGCGCTGTCCAAAGTTCAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 955 AAGCGCGCTGTCCAAAGTTCAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
 841 GTCATAGCTTACTGTCTATTCGCGGGCGCGCACTGGGCTGGGAGAGGTCAAAGTTCTAC 900
 1015 GTCATAGCTTACTGTCTATTCGCGGGCGCGCACTGGGCTGGGAGAGGTCAAAGTTCTAC 1074
 901 ATCGAGCTCAAGAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGCTCTCAG 960
 1075 ATCGAGCTCAAGAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGCTCTCAG 1134
 961 ATCTCTCTGCTTTCGCTGGAGGCAATATGACAGCCACGAGGAGTTCCACTGCATGATG 1020
 1135 ATCTCTCTGCTTTCGCTGGAGGCAATATGACAGCCACGAGGAGTTCCACTGCATGATG 1194
 1021 AACTGGACGACATCGAGCTTCGCTGCTCAAGGGGCTCTTCGCGCTCGCTGCGCTACCTC 1080
 1195 AACTGGACGACATCGAGCTTCGCTGCTCAAGGGGCTCTTCGCGCTCGCTGCGCTACCTC 1254
 1081 ACTTGGGGCGAGAGACCAAGAGGTTCATACGGATAAAGCTGCGCGCGGCTCCATCCGGGCC 1140
 1255 ACTTGGGGCGAGAGACCAAGAGGTTCATACGGATAAAGCTGCGCGCGGCTCCATCCGGGCC 1314
 1141 GTGGTCAACATCTTTCTGGTGGCAGGCGCTGTGCTCTATCTCTGCTGCGCATTTCTTTGCC 1200
 1315 GTGGTCAACATCTTTCTGGTGGCAGGCGCTGTGCTCTATCTCTGCTGCGCATTTCTTTGCC 1374
 1201 GCTGTGAGGTGCTGGAGAGTTCGCTTTTCAAGGAGGAGCGCGCGCTTTTTCGCGGCC 1260
 1375 GCTGTGAGGTGCTGGAGAGTTCGCTTTTCAAGGAGGAGCGCGCGCTTTTTCGCGGCC 1434

Db 469 CAGGCCAAGATGAGCGCATGTTCCAGGATGGGTTTCAGCGCGCCACGATGAGGAG 528
Qy 121 GCGTGGGCTTCGCGCATTTGCGACACCTCGACTTTGAGCAGCCGCGGCTTCGAGATG 180
Db 529 GCGTGGGCTTCGCGCATTTGCGACACCTCGACTTTGAGCAGCCGCGGCTTCGAGATG 588
Qy 181 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGGAGAGGGGCTGAAAGCGCCGCTCGAG 240
Db 589 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGGAGAGGGGCTGAAAGCGCCGCTCGAG 648
Qy 241 GGAGACATCCATTATCAGCGAGCGAGAGCTCTCTGCGGCCCTCGGGCTCCAAGAC 300
Db 649 GGAGACATCCATTATCAGCGAGCGAGAGCTCTCTGCGGCCCTCGGGCTCCAAGAC 708
Qy 301 CAGTGGGAGGTGGTGGGAAATTCGGGGCCGAGGAGAGGGGCTGAAAGCGCCGCTCGAG 360
Db 709 CAGTGGGAGGTGGTGGGAAATTCGGGGCCGAGGAGAGGGGCTGAAAGCGCCGCTCGAG 768
Qy 361 GCAGGCTGAAAGCTGACCAAGCCATCAGGAGCATGTTGCTGCTGGGCTACCTACGCC 420
Db 769 GCAGGCTGAAAGCTGACCAAGCCATCAGGAGCATGTTGCTGCTGGGCTACCTACGCC 828
Qy 421 ATCTCTGACGGCGCTTACCTGGGCTGTTTCTCATCATCTTCGCGCCGCTTGTGTCTGC 480
Db 829 ATCTCTGACGGCGCTTACCTGGGCTGTTTCTCATCATCTTCGCGCCGCTTGTGTCTGC 888
Qy 481 TACACGGCAAGATCCTCATCGCTGCTGTACGAGGAGAAATGAAGACGGCGAGTGGTG 948
Db 889 TACACGGCAAGATCCTCATCGCTGCTGTACGAGGAGAAATGAAGACGGCGAGTGGTG 948
Qy 541 CCGCTGCGGAGCTCGTACGAGCCATAGCAAGCGCTGCGCGCCCGGCTTCCCAAG 600
Db 949 CCGCTGCGGAGCTCGTACGAGCCATAGCAAGCGCTGCGCGCCCGGCTTCCCAAG 1008
Qy 601 CTGGGCGCGGAGTGGTGAACCTAGCGAGATCATCGAGCTGCTGATGACGTCATCTCG 660
Db 1009 CTGGGCGCGGAGTGGTGAACCTAGCGAGATCATCGAGCTGCTGATGACGTCATCTCG 1068
Qy 661 TAGCTGGTGGTGGGCAACCTCATGTACACAGCTTCCGGGCTGCGCGCTGTCGAG 720
Db 1069 TAGCTGGTGGTGGGCAACCTCATGTACACAGCTTCCGGGCTGCGCGCTGTCGAG 1128
Qy 721 AAGTCTGCTGCAATATCGCACGGCGTGTGCTGCTGCGCTTCCCTTAAAGAACCTC 780
Db 1129 AAGTCTGCTGCAATATCGCACGGCGTGTGCTGCTGCGCTTCCCTTAAAGAACCTC 1188
Qy 781 AAGGCGCTGTCCAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 1189 AAGGCGCTGTCCAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
Qy 841 GTCATAGCTACTGCTATCGCGGGCGCGGACTGCGGCTGCGGAGAGTCAAGTCTAC 900
Db 1249 GTCATAGCTACTGCTATCGCGGGCGCGGACTGCGGCTGCGGAGAGTCAAGTCTAC 1308
Qy 901 ATCGAGCTCAAGAGTTCCCATCTCCATTTGGCATCATCTGCTGCTGCTGCTGCTGCTG 960
Db 1309 ATCGAGCTCAAGAGTTCCCATCTCCATTTGGCATCATCTGCTGCTGCTGCTGCTGCTG 1368
Qy 961 ATCTTCTGCTTCCGCTGAGGAGCAATATGACAGCAGCCGAGGAGTTCACCTGATGATG 1020
Db 1369 ATCTTCTGCTTCCGCTGAGGAGCAATATGACAGCAGCCGAGGAGTTCACCTGATGATG 1428
Qy 1021 AACTGACGACATCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1429 AACTGACGACATCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1488
Qy 1081 ACCTGGGCGAGCAGACAGGAGGCTCATCAGCGATAAAGCTGCGGCTTCCATCCGCGCC 1140
Db 1489 ACCTGGGCGAGCAGACAGGAGGCTCATCAGCGATAAAGCTGCGGCTTCCATCCGCGCC 1548
Qy 1141 GTGGTCAACATCTTCTGCTGGGCAAGGCGCTGTTGCTCTATCTCTGCTGCTGCTGCTGCTG 1200

Db 1549 GTGTCACATCTTTCTGTGGTGGCAAGCGGCTGTGTCTATATCTCTGCTGCTTCTTGGCC 1608
Qy 1201 GCTGTGAGGTGCTGGAGAAAGTCGCTCTTTCAGGAAAGCAGCGCGCTTTTCCCGGCC 1260
Db 1609 GCTGTGAGGTGCTGGAGAAAGTCGCTCTTTCAGGAAAGCAGCGCGCTTTTCCCGGCC 1668
Qy 1261 TGCTACAGCGCGCAGCGGCGCTTGAAGTCTTGGGGGTGACGCTGCGCTGCGGCTCGTC 1320
Db 1669 TGCTACAGCGCGCAGCGGCGCTTGAAGTCTTGGGGGTGACGCTGCGCTGCGGCTCGTC 1728
Qy 1321 GTCTTACAGCTGCTATGCGCCATTTATGTGCGGCACATTCGCGCTGCTCATGGGCTCAC 1380
Db 1729 GTCTTACAGCTGCTATGCGCCATTTATGTGCGGCACATTCGCGCTGCTCATGGGCTCAC 1788
Qy 1381 GCGAGCTCAGCGCGCGCGCTCTGTCTTCTGCTGCGCAGCTTTCACCTGCGGCTG 1440
Db 1789 GCGAGCTCAGCGCGCGCGCTCTGTCTTCTGCTGCGCAGCTTTCACCTGCGGCTG 1848
Qy 1441 CTCTGGGCAAGCTGCTGTGGCACAAGTCTTCTTGGAGCTGCGCATCTTGTCTATCGGC 1500
Db 1849 CTCTGGGCAAGCTGCTGTGGCACAAGTCTTCTTGGAGCTGCGCATCTTGTCTATCGGC 1908
Qy 1501 GCGATCTGACGCTGCTGCGGCTTGTGACCTCCCTCGAGGCGCTCATCGAGCTTACCGA 1560
Db 1909 GCGATCTGACGCTGCTGCGGCTTGTGACCTCCCTCGAGGCGCTCATCGAGCTTACCGA 1968
Qy 1561 ACCAAGCGGAGGACTAG 1578
Db 1969 ACCAAGCGGAGGACTAG 1986

RESULT 4

BC053582

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

2519 bp mRNA linear PRI 30-JUN-2004
Homo sapiens vesicular inhibitory amino acid transporter, mRNA
(CDNA clone MGC:61677 IMAGE:5180816), complete cds.

BC053582

BC053582.1

GI:31566391

MGC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2519)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

McKernan, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Pahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Small, D.E.,

Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2519)

Strausberg, R.

Direct Submission

Submitted (09-JUN-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

Query Match		100.0%;	Score 1578;	DB 9;	Length 1592;		
Best Local Similarity		100.0%;	Mismatches 0;	Indels 0;	Gaps 0;		
Matches 1578;		Conservative 0;	Indels 0;	Gaps 0;			
QY	1	ATGGCCACCTTGTCTCGCAGCAAGCTGTCTCAAGCTGGCCACGTCGTCGTGTCTCAACAAAGTCC	60				
DB	6	ATGGCCACCTTGTCTCGCAGCAAGCTGTCTCAAGCTGGCCACGTCGTCGTGTCTCAACAAAGTCC	65				
QY	61	CAGGCCAAGATGAGCGGCATGTTCCGCCAGGATGGGTTTTCAGGGCGGCCACGGATGAGGAG	120				
DB	66	CAGGCCAAGATGAGCGGCATGTTCCGCCAGGATGGGTTTTCAGGGCGGCCACGGATGAGGAG	125				
QY	121	CGCGTGGGCTTCGGGCATTTGCGACGACCTCGACTTTTGAGCACCGCCAGGGGCTCGCAGATG	180				
DB	126	CGCGTGGGCTTCGGGCATTTGCGACGACCTCGACTTTTGAGCACCGCCAGGGGCTCGCAGATG	185				
QY	181	GACATCTGAAAGCGGGAGAGCCCTGCGGGGACGAGGCGCTGAAAGCCCGCTCGAG	240				
DB	186	GACATCTGAAAGCGGGAGAGCCCTGCGGGGACGAGGCGCTGAAAGCCCGCTCGAG	245				
QY	241	GGAGACATCCATTATCAGCGAGGCGAGGAGCTCTCTGCGCGCCCTCCGGCTCCAAGGAC	300				
DB	246	GGAGACATCCATTATCAGCGAGGCGAGGAGCTCTCTGCGCGCCCTCCGGCTCCAAGGAC	305				
QY	301	CAGGTGGGAGTGTGGCGAATTCGGGGGCCACGACAAAGCCCAAATACCGCGCTGGGAG	360				
DB	306	CAGGTGGGAGTGTGGCGAATTCGGGGGCCACGACAAAGCCCAAATACCGCGCTGGGAG	365				
QY	361	GCAGGCTGGAAAGCTGACCAAGCCATCCAGGGCATGTCGTGCTGGGCTTACCCCTACGCC	420				
DB	366	GCAGGCTGGAAAGCTGACCAAGCCATCCAGGGCATGTCGTGCTGGGCTTACCCCTACGCC	425				
QY	421	ATCTCTGACGCGCGCTACCTGGGGTGTCTTCATCATCTTCGCGCGCGTGTGTGTCTGTCG	480				
DB	426	ATCTCTGACGCGCGCTACCTGGGGTGTCTTCATCATCTTCGCGCGCGTGTGTGTCTGTCG	485				
QY	481	TACACCGGCAAGATCTCTCATCGCTGCTGTATCGAGGAGATGAAGACGGCGAGTGGTG	540				
DB	486	TACACCGGCAAGATCTCTCATCGCTGCTGTATCGAGGAGATGAAGACGGCGAGTGGTG	545				
QY	541	CGGTGCGGGACTCGTACGTGGCCATAGCCCAAGCCCTGCTGCGCGCCCGCGCTTCCCAACG	600				
DB	546	CGGTGCGGGACTCGTACGTGGCCATAGCCCAAGCCCTGCTGCGCGCCCGCGCTTCCCAACG	605				
QY	601	CTGGGCGGCGAGTGGTGAAGCTAGCGCAGATCATTCGAGCTGTGTGATGACGTGCATCTGT	660				
DB	606	CTGGGCGGCGAGTGGTGAAGCTAGCGCAGATCATTCGAGCTGTGTGATGACGTGCATCTGT	665				
QY	661	TACGTGTGTGAGTGGCAACTCATGTACAAAGCTTCCCGGGGCTGCGCGTGTGCGAG	720				
DB	666	TACGTGTGTGAGTGGCAACTCATGTACAAAGCTTCCCGGGGCTGCGCGTGTGCGAG	725				
QY	721	AAGTCTGTGTCATTATCGCCACGGCGGTGCTGTGCTTGGCGCTTCTTTAAGAACTCT	780				
DB	726	AAGTCTGTGTCATTATCGCCACGGCGGTGCTGTGCTTGGCGCTTCTTTAAGAACTCT	785				
QY	781	AAGCCGTGTCCAAGTTCAGTCTGTGCACTCTGCGCCACATTCGTCAATATCATCTGT	840				
DB	786	AAGCCGTGTCCAAGTTCAGTCTGTGCACTCTGCGCCACATTCGTCAATATCATCTGT	845				
QY	841	GTCAATAGCTACTGTCTATCGCGGCGCGGACTTGGGCGCTGGGAGAGGTCAAGTTCATC	900				
DB	846	GTCAATAGCTACTGTCTATCGCGGCGCGGACTTGGGCGCTGGGAGAGGTCAAGTTCATC	905				
QY	901	ATCGAGCTCAAGAGTTCCTCCATCTCCATTGGCATCATCGTGTTCAGCTACAGTCTCAG	960				
DB	906	ATCGAGCTCAAGAGTTCCTCCATCTCCATTGGCATCATCGTGTTCAGCTACAGTCTCAG	965				
QY	961	ATCTTCTGTGCTTCTGCTGGAGGGCAATATGACAGACGCCACGAGTTCCTATGATGATG	1020				
DB	966	ATCTTCTGTGCTTCTGCTGGAGGGCAATATGACAGACGCCACGAGTTCCTATGATGATG	1025				
QY	1021	AACTGGAGCGACATCGCAGGCTCGTGTCTCAAGGGGCTCTTCGCGGCTCGTTCGCTACCTC	1080				

Db	1026																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												</
----	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

181	Db	GACATCCTGAAACCCGAGGAGAGCCCTCGGGGACGAGGGCGCTGTAAGCGCCGCTCGAG	240
241	Qy	GGAGACATCCATTTATCAGCGAGGACGAGCTCTCTGCGCCCTTCGGCTCCAAGGAC	300
241	Db	GGAGACATCCATTTATCAGCGAGGACGAGCTCTCTGCGCCCTTCGGCTCCAAGGAC	300
301	Qy	CAGGTGGAGGTTGGTGGCNAATTCGGGGGCGACGAAAGCCNAATACAGGCGTGGAG	360
301	Db	CAGGTGGAGGTTGGTGGCNAATTCGGGGGCGACGAAAGCCNAATACAGGCGTGGAG	360
361	Qy	GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTCGTGCTGGGCTACCCCTACGCC	420
361	Db	GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTCGTGCTGGGCTACCCCTACGCC	420
421	Qy	ATCCTGCACCGCGCTACCTCGGGGTTGTTTCTCATCATCTTCGCCCGCTGTGTGCTGC	480
421	Db	ATCCTGCACCGCGCTACCTCGGGGTTGTTTCTCATCATCTTCGCCCGCTGTGTGCTGC	480
481	Qy	TACACCGGCNAAGATCCTCATTCGCGTGCTGTGACGAGGAAATGAAGACGGCGAGGTGGT	540
481	Db	TACACCGGCNAAGATCCTCATTCGCGTGCTGTGACGAGGAAATGAAGACGGCGAGGTGGT	540
541	Qy	CGCGTGGGACGTCGTACGTGGGCATAGCAACGCCCTGCTGCGCCCGCGCTCCCAAG	600
541	Db	CGCGTGGGACGTCGTACGTGGGCATAGCAACGCCCTGCTGCGCCCGCGCTCCCAAG	600
601	Qy	CTGGGCGGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGTGATGACGTGCATCTCG	660
601	Db	CTGGGCGGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGTGATGACGTGCATCTCG	660
661	Qy	TACGTGGTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGCTGCCGTTGTCGAG	720
661	Db	TACGTGGTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGCTGCCGTTGTCGAG	720
721	Qy	AAGTCTGTGTCATATACGCCACGGCGTGCTGCTGCTGCGCTTCTCTTAGAACCTC	780
721	Db	AAGTCTGTGTCATATACGCCACGGCGTGCTGCTGCTGCGCTTCTCTTAGAACCTC	780
781	Qy	AAGCGCGTGCCTCAAGTTTCAGTCTGTGTCACCTTCGCCACATCTCGTCAATCAATCTCG	840
781	Db	AAGCGCGTGCCTCAAGTTTCAGTCTGTGTCACCTTCGCCACATCTCGTCAATCAATCTCG	840
841	Qy	GTCAATACCTACTGTCTATACGCGGGCGCGACTGGGCTTGGAGAGGTCAGTCTTAC	900
841	Db	GTCAATACCTACTGTCTATACGCGGGCGCGACTGGGCTTGGAGAGGTCAGTCTTAC	900
901	Qy	ATCGAGTCAAGAGTTCCCATCTCAATTTGGCATCATCGTGTTCAGTACACGTCTCAG	960
901	Db	ATCGAGTCAAGAGTTCCCATCTCAATTTGGCATCATCGTGTTCAGTACACGTCTCAG	960
961	Qy	ATCTTCCTGCTTCGTGGAGGGCAATATGACGACGCCAGCGAGTTCCACTGCATGATG	1020
961	Db	ATCTTCCTGCTTCGTGGAGGGCAATATGACGACGCCAGCGAGTTCCACTGCATGATG	1020
1021	Qy	AACTTGACGACATTCGAGCTCGTGTCTCAAGGGCTCTTCGCGCTCGTCCCTACCTC	1080
1021	Db	AACTTGACGACATTCGAGCTCGTGTCTCAAGGGCTCTTCGCGCTCGTCCCTACCTC	1080
1081	Qy	ACCTGGGCGACGAGACAAGAGGATCATACCGATTAACCTGCGCGCTCCATCCGGGCC	1140
1081	Db	ACCTGGGCGGACGAGACAAGAGGATCATACCGATTAACCTGCGCGCTCCATCCGGGCC	1140
1141	Qy	GTGGTCAACATCTTTCTGTGGTGGCCAAAGCGCTGTTCCTATCTCTGCGCATCTTTGCC	1200
1141	Db	GTGGTCAACATCTTTCTGTGGTGGCCAAAGCGCTGTTCCTATCTCTGCGCATCTTTGCC	1200
1201	Qy	GCTGTCGAGGTGCTGGAGAAATCGCTCTTCCAGGAAGCGACCGCGCTTTTTCCTCGGCC	1260
1201	Db	GCTGTCGAGGTGCTGGAGAAATCGCTCTTCCAGGAAGCGACCGCGCTTTTTCCTCGGCC	1260
1261	Qy	TGCTACAGCGCGACGCGCGCTGAAATCTCTGGGGCTGACGTGCGCTCGGCTCGCTCGT	1320
1261	Db	TGCTACAGCGCGACGCGCGCTGAAATCTCTGGGGCTGACGTGCGCTCGGCTCGCTCGT	1320

Qy	1321	GTCTTACGCTGCTCATGGCCATTTATGTGCCGCACTTTCGCGTCTCATGGCCTCACC	1380
Db	1321	GTCTTACGCTGCTCATGGCCATTTATGTGCCGCACTTTCGCGTCTCATGGCCTCACC	1380
Qy	1381	GGCAGCCTCACGGCGCGCGCTCTGTTTCTTGCTGCGCAGCCTCTTTTACCTGCGCCTG	1440
Db	1381	GGCAGCCTCACGGCGCGCGCTCTGTTTCTTGCTGCGCAGCCTCTTTTACCTGCGCCTG	1440
Qy	1441	CTCTGGCGCAAGCTGTGTGGCAACCAAGTCTTCTTTCGAGTTCGCCATCTTTCGTCATCGGC	1500
Db	1441	CTCTGGCGCAAGCTGTGTGGCAACCAAGTCTTCTTTCGAGTTCGCCATCTTTCGTCATCGGC	1500
Qy	1501	GGCATCTGCAGCTGTCCGCTTCGTGCACTCCCTCGAGGGCCTCATCGAAGCCTACCGA	1560
Db	1501	GGCATCTGCAGCTGTCCGCTTCGTGCACTCCCTCGAGGGCCTCATCGAAGCCTACCGA	1560
Qy	1561	ACCAACGCGGAGCTAG	1578
Db	1561	ACCAACGCGGAGCTAG	1578

RESULT 2

AY044836

LOCUS

DEFINITION Homo sapiens vesicular inhibitory amino acid transporter mRNA, complete cds.

ACCESSION AY044836

VERSION AY044836.1

KEYWORDS GI:17975776

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 1592)

CHESLER,S.D., SIMONSON,W.T., SWEET,I.R. and HAMMERLE,L.P.

EXPRESSION OF THE VESICULAR INHIBITORY AMINO ACID TRANSPORTER IN PANCREATIC ISLET CELLS: DISTRIBUTION OF THE TRANSPORTER WITHIN RAT ISLETS

JOURNAL Diabetes

MEDLINE 51 (6), 1763-1771 (2002)

PUBMED 12027589

PMID 12031963

REFERENCE

2 (bases 1 to 1592)

CHESLER,S.D. and GOHLKE,P.R.

DIRECT SUBMISSION

TITLE Submitted (10-JUL-2001) Medicine, University of Washington, UW Health Sciences Center, Room K-161, Box 357710, Seattle, WA 98195-7710, USA

FEATURES

source

1..1592

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="20"

/tissue_type="brain"

6..1583

/note="VIAAT; vesicular neurotransmitter transporter"

/codon_start=1

/product="vesicular inhibitory amino acid transporter"

/protein_id="AAK98782.1"

/db_xref="GI:17975776"

/translation="MATLLSKLSNVATSVNSKQAKNSGFMARMGFOAATDEAVGFGV
AHCDDLFQHRQGLQMDILKAEGEPQDEGAAPVEGDHYRGSGAPLPSPGSKQV
GGGGEFGCHDKPKTIAWEAGANNVTNAIQGMVLGPLYAILHGGYLGFLFIIFAAVCC
YTKILIALCYEENEDVEVVRDYSVYAIANACAPRFTLGGRVNVVAQIIELVMTFC
ILYVVSGLMYNSFGCLPVYSQKWSIATATVLLPCAFKLNKLVSKFSLCTLAHFV
INILVIAYCISLRADWAEVKFYIDVKKFPISIGIIVFVTSQIFLPSLEGNNQKPS
EFHCCMWNTHAACVCLKGLFALVAITWADETREKVIDTNLPGSRATVNVIFLVAKALL
SYLPFFAAVVEUEKSLFQBSGRAFFPACISGGGRKLSKSGLTURCALVWFTLLMAITV
PHFALLMGLTGSITGLTGKFLPLSLFHLRLWKLWHQVFFDVAFIVIGICSVSGFV
VHSLEGLIARVTRWAE"

CDS

1..1583

/note="VIAAT; vesicular neurotransmitter transporter"

/codon_start=1

/product="vesicular inhibitory amino acid transporter"

/protein_id="AAK98782.1"

/db_xref="GI:17975776"

/translation="MATLLSKLSNVATSVNSKQAKNSGFMARMGFOAATDEAVGFGV
AHCDDLFQHRQGLQMDILKAEGEPQDEGAAPVEGDHYRGSGAPLPSPGSKQV
GGGGEFGCHDKPKTIAWEAGANNVTNAIQGMVLGPLYAILHGGYLGFLFIIFAAVCC
YTKILIALCYEENEDVEVVRDYSVYAIANACAPRFTLGGRVNVVAQIIELVMTFC
ILYVVSGLMYNSFGCLPVYSQKWSIATATVLLPCAFKLNKLVSKFSLCTLAHFV
INILVIAYCISLRADWAEVKFYIDVKKFPISIGIIVFVTSQIFLPSLEGNNQKPS
EFHCCMWNTHAACVCLKGLFALVAITWADETREKVIDTNLPGSRATVNVIFLVAKALL
SYLPFFAAVVEUEKSLFQBSGRAFFPACISGGGRKLSKSGLTURCALVWFTLLMAITV
PHFALLMGLTGSITGLTGKFLPLSLFHLRLWKLWHQVFFDVAFIVIGICSVSGFV
VHSLEGLIARVTRWAE"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 18:32:15 ; Search time 6606 Seconds
(without alignments)
11296.268 Million cell updates/sec

Title: US-09-940-919-1
Perfect score: 1578
Sequence: 1 atggccaccttgcgcag,.....gaaccaacgaggactag 1578

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_btg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

ALIGNMENTS

RESULT 1
AX398633
LOCUS AX398633 1578 bp DNA linear PAT 27-MAY-2002
DEFINITION Sequence 1 from Patent WO0220753.
ACCESSION AX398633
VERSION AX398633.1 GI:21261222
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Friddle,C.J., Gerhardt,B. and Hu,Y.
TITLE Human gaba transporter protein and polynucleotides encoding the same
JOURNAL Patent: WO 0220753-A 1 14-MAR-2002;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..1578
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	6	AX398633 Sequence
2	1578	100.0	1592	6	AX044836 Homo sapi
3	1578	100.0	2261	6	AX398635 Sequence
4	1578	100.0	2519	9	BC053582 Homo sapi
5	1578	100.0	2550	9	AK055051 Homo sapi
6	1576.4	99.9	1763	6	AX497262 Sequence
7	1576.4	99.9	1800	6	AX259431 Sequence
8	1574.8	99.8	1734	6	AX497264 Sequence
9	1574.8	99.8	1734	6	AX497268 Sequence
10	1574.8	99.8	1734	6	AX497270 Sequence
11	1574.8	99.8	1734	6	AX497272 Sequence
12	1567.4	99.3	1733	6	AX497266 Sequence
13	1518.8	96.2	2756	9	AB062931 Macaca fa
14	1424.8	90.3	2174	9	BC036458 Homo sapi
15	1358	86.1	1473	6	CQ727095 Sequence
16	1352.2	85.7	2392	6	AR438018 Sequence
17	1352.2	85.7	2392	10	AF030253 Rattus no
18	1345.8	85.3	2814	10	BC052020 Mus muscu
19	1344.2	85.2	1648	6	AX147507 Sequence